07-23-0 JC03 Rec'd PCT/PTO 20 JUL 2001

PTO-1390 (REV 12-29-99)	U S DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE	ATTORNEY'S DOCKET NUMBER MWH-008US

TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371

U.S APPLICATION NO (If known, see 37 CFR 1 5): 09/889866

INTERNATIONAL APPLICATION NO.	INTERNATIONAL FILING DATE	PRIORITY DATE CLAIMED
PCT/US00/22175	11 August 2000 (11.08.00)	24 August 1999 (24.08.99)

TITLE OF INVENTION: Drug Target Isogenes: Polymorphisms in the Immunoglobulin E Receptor Beta Chain Gene

APPLICANT(S) FOR DO/EO/US

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

1. [X] This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
2. [ ] This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
3. [X] This express request to begin national examination procedures (35 U S.C. 371(f)) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT Articles 22 and 39(1).
4. [X] A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
5. [X] A copy of the International Application as filed (35 U.S.C. 371(c)(2))
   a. [ ] is transmitted herewith (required only if not transmitted by the International Bureau).
   b. [ ] has been transmitted by the International Bureau.
   c. [X] is not required, as the application was filed in the United States Receiving Office (RO/US).
6. [ ] A translation of the International Application into English (35 U.S.C. 371(c)(2)).
7. [X] Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3))
   a. [ ] are transmitted herewith (required only if not transmitted by the International Bureau).
   b. [ ] have been transmitted by the International Bureau.
   c. [ ] have not been made; however, the time limit for making such amendments has NOT expired.
   d. [X] have not been made and will not be made.
8. [ ] A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).
9. [X] An oath or declaration of the inventor(s) (35 U.S.C. 371(c)(4)).
10. [ ] A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).

Items 11. to 16. below concern document(s) or information included:

11. [X] An Information Disclosure Statement under 37 CFR 1.97 and 1.98.
12. [ ] An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.
13. [ ] A FIRST preliminary amendment.
    [ ] A SECOND or SUBSEQUENT preliminary amendment.
14. [ ] A substitute specification.
15. [ ] A change of power of attorney and/or address letter.
16. [X] Other items or information:

The claims in the International Application were modified by a PCT Article 34 Amendment. A copy of this Amendment is enclosed herein. It is Applicant's intent that the national stage of processing be performed on the same claims in the International Application as amended by the Article 34 Amendment.

JC17 Rec'd PCT/PTO 20 JUL 2001

U.S. APPLICATION NO (if known, see 37 CFR 1.5)	INTERNATIONAL APPLICATION NO	ATTORNEY'S DOCKET NUMBER
09/889866	PCT/US00/22175	MWH-008US

17. [X] The following fees are submitted:

BASIC NATIONAL FEE (37 CFR 1.492 (a) (1) - (5)) :

Neither international preliminary examination fee (37 CFR 1.482) nor international search fee (37 CFR 1.445(a)(2)) paid to USPTO and International Search Report not prepared by the EPO or JPO **$970.00**

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but International Search Report prepared by the EPO or JPO **$840.00**

International preliminary examination fee (37 CFR 1.482) not paid to USPTO but international search fee (37 CFR 1.445(a)(2)) paid to USPTO **$690.00**

International preliminary examination fee paid to USPTO (37 CFR 1.482) but all claims did not satisfy provisions of PCT Article 33(1)-(4) **$670.00**

International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(1)-(4) **$96.00**

	CALCULATIONS PTO USE ONLY
ENTER APPROPRIATE BASIC FEE AMOUNT =	$ 96.00
Surcharge of **$130.00** for furnishing the oath or declaration later than [ ] 20 [ ] 30 months from the earliest claimed priority date (37 CFR 1.492(e)).	$ 0.00

CLAIMS	NUMBER FILED	NUMBER EXTRA	RATE	
Total claims	27 - 20 =	7	X **$18.00**	$ 126.00
Independent claims	11 - 3 =	8	X **$78.00**	$ 640.00
MULTIPLE DEPENDENT CLAIM(S) (if applicable)			+ **$260.00**	$ 0.00
TOTAL OF ABOVE CALCULATIONS =				$ 866.00
Reduction of 1/2 for filing by small entity, if applicable. A Small Entity Statement must also by filed (Note 37 CFR 1.9, 1.27, 1.28).				$ 0.00
SUBTOTAL =				$ 866.00
Processing fee of **$130.00** for furnishing the English translation later than [ ]20 [ ]30 months from the earliest claimed priority date (37 CFR 1.492(f)). +				$ 0.00
TOTAL NATIONAL FEE =				$ 866.00
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). **$40.00** per property +				$ 0.00
TOTAL FEES ENCLOSED =				$ 866.00
				Amount to be refunded: $
				charged: $

a. [ ] A check in the amount of $_____ to cover the above fees is enclosed.

b. [X] Please charge my Deposit Account No. 50-1293 in the amount of $866.00 to cover the above fees. A duplicate copy of this sheet is enclosed.

c. [X] The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 50-1293. A duplicate copy of this sheet is enclosed.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Inna Shtivelband
Genaissance Pharmaceuticals, Inc.
Five Science Park
New Haven, CT 06511

SIGNATURE

Inna Shtivelband
NAME

44,337
REGISTRATION NUMBER

Practitioner's Docket No. MWH-008US *PATENT*

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:
Genaissance Pharmaceuticals, Inc.

Examiner:
Deborah Crouch

Application No.: To be assigned
Filed: July 20, 2001

Group No.: 1632

For: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE IMMUNOGLOBULIN E RECEPTOR BETA CHAIN GENE

Assistant Commissioner for Patents
Washington, D.C. 20231

REQUEST FOR EXPEDITED EXAMINATION OF NATIONAL STAGE APPLICATION PURSUANT TO 35 U.S.C. 371(f) and 37 C.F.R. 1.496)

Applicants believe that they have met the requirements of 35 U.S.C. 371(f) and 37 C.F.R. 1.496(b) by submitting the following herein:

(a) **Application filed under 35 U.S.C. 371(c).**

Applicants herein file a National Application pursuant to 35 U.S.C. 371. The claims in this Application are identical to the claims in International Application PCT/US00/22175 as amended by a PCT Article 34 Amendment, a copy of which is enclosed herein. The United States Patent and Trademark Office indicated in an International Preliminary Examination Report (copy enclosed herein) that the claims in the International Application, as amended, satisfy the criteria of PCT Articles 33(1)-(4) as to novelty, inventive step and industrial applicability:

(b) **Basic National Fee**

The Basic National Fee pursuant to 37 C.F.R. 1.492(a)(4) is enclosed.

Respectfully submitted,

Reg. No. 44,337
Tel. No. 203-786-3529

Inna Shtivelband
Genaissance Pharmaceuticals, Inc.
Five Science Park
New Haven, CT 06511

PATENT COOPERATION TREATY

IN THE UNITED STATES RECEIVING OFFICE

In Re: International Application of: Genaissance Pharmaceuticals, Inc.

International Application No.: PCT/US00/22175
Attorney Docket No.: MWH-0008PCT
Authorized Officer: Taieb Akremi
International Filing Date: 11 August 2000

For: Drug Target Isogenes: Polymorphisms in the Immunoglobulin E Receptor Beta ChainGene

Box PCT
Assistant Commissioner for Patents
Washington, DC 20231
ATTENTION: IPEA/US

AMENDMENT TO CLAIMS UNDER ARTICLE 34(2)(b)

This paper is filed with the Demand to focus the claimed invention on isogenes of the Immunoglobulin E Receptor Beta Chain Gene that occur in the general population and detection of haplotypes that define these isogenes. The differences between the claims as filed and amended are shown on the attached copies of the claims pages with underlining indicating added material and brackets indicating deleted material. A summary of the claim amendments and support therefor is set forth below:

Claim Number	Action	Support
1	Amended	p. 6, lines 24-29and p. 13, lines 22-24
2	Cancelled	
3-5	Unchanged	
6	Amended	p. 6, lines 24-29 and p. 13, lines 22-24
7	Amended	p. 13, lines 22-24; p. 15, lines 2-12; and Figure 2
8-9	Unchanged	
10	Amended	p. 16, lines 31-32
11-16	Unchanged	
17	Amended	p. 25, lines 21-25 and lines 28-30
18	Amended	Claim 17, as filed
19	Amended	p. 26 lines 2-6 and line 15
20	Amended	Claim 19, as filed

Claim Number	Action	Support
21	Amended	p. 26, lines 16-20 and p. 28, line 36 to p. 29, line 6
22	Amended	claims 22-23, as filed; p. 30, line 37 to p. 31, line 2; p. 31, line 36 to p. 32, line 2; and p. 32, lines 29-32
23	Cancelled	
24	Amended	Claims 22-24, as filed
25-26	Unchanged	
27-29	New	p. 6, lines 28-32; p. 40, lines 12-13; p. 26, lines 16-20, and p. 39, lines 29-31

Replacement pages 42-45 are attached herein. Applicants respectfully request acceptance of these claim amendments and that the International Preliminary Examination be based thereon. If any questions arise regarding this submission, please contact the undersigned attorney at the phone number below.

Respectfully submitted,

Sandra L. Shaner

Reg. Num. P-47,934
Tel. No. 203-786-3468

Sandra L. Shaner
Genaissance Pharmaceuticals, Inc.
Five Science Park
New Haven, CT 06511

March 26, 2001

What is Claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) a first nucleotide sequence which is a polymorphic variant of a reference sequence for immunoglobulin E receptor beta chain (IGERB) gene or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:1, and the polymorphic variant comprises an IGERB isogene defined by a haplotype selected from the group consisting of haplotypes 1-12 in Table 5[at least one polymorphism selected from the group consisting of adenine at PS2, thymine at PS3, cytosine at PS4, cytosine at PS5, thymine at PS6, guanine at PS7, guanine at PS8, guanine at PS9, adenine at PS10, cytosine at PS12 and thymine at PS13]; and

(b) a second nucleotide sequence which is complementary to the first nucleotide sequence.

[2. The isolated polynucleotide of claim 1 which comprises an IGERB isogene.]

3. The isolated polynucleotide of claim 1 which is a DNA molecule and comprises both the first and second nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.

4. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 1, wherein the organism expresses an IGERB protein encoded by the first nucleotide sequence.

5. The recombinant organism of claim 4 which is a nonhuman transgenic animal.

6. The isolated polynucleotide of claim 1, wherein the first nucleotide sequence is a polymorphic variant of a fragment of the IGERB isogene[gene], the fragment comprising one or more polymorphisms selected from the group consisting of adenine at PS2, thymine at PS3, cytosine at PS4, cytosine at PS5, thymine at PS6, guanine at PS7, guanine at PS8, guanine at PS9, adenine at PS10, cytosine at PS12 and thymine at PS13.

7. An isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the IGERB cDNA or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:2 and the polymorphic variant comprises a coding sequence of an IGERB isogene defined by a haplotype selected from the group consisting of haplotypes 1-12 in Table 5[at least one polymorphism selected from the group consisting of thymine at a position corresponding to nucleotide 132, thymine at a position corresponding to nucleotide 428, guanine at a position corresponding to nucleotide 473, and guanine at a position corresponding to nucleotide 532].

8. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 7, wherein the organism expresses an immunoglobulin E receptor beta chain (IGERB) protein encoded by the polymorphic variant sequence.

9. The recombinant organism of claim 8 which is a nonhuman transgenic animal.

10. An isolated polypeptide comprising an amino acid sequence which is a polymorphic variant of a reference sequence for the IGERB protein or a fragment thereof, wherein the reference sequence comprises SEQ ID NO: 3 and the polymorphic variant is encoded by an isogene defined by one of

the haplotypes shown in Table 5.[comprises one or more variant amino acids selected from the group consisting of methionine at a position corresponding to amino acid position 143, cysteine at a position corresponding to amino acid position 158, and alanine at a position corresponding to amino acid position 178 .]

11. An isolated antibody specific for and immunoreactive with the isolated polypeptide of claim 10.

12. A method for screening for drugs targeting the isolated polypeptide of claim 10 which comprises contacting the IGERB polymorphic variant with a candidate agent and assaying for binding activity.

13. A composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in the immunoglobulin E receptor beta chain (IGERB) gene at a polymorphic site selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.

14. The composition of claim 13, wherein the genotyping oligonucleotide is an allele-specific oligonucleotide that specifically hybridizes to an allele of the IGERB gene at a region containing the polymorphic site.

15. The composition of claim 14, wherein the allele-specific oligonucleotide comprises a nucleotide sequence selected from the group consisting of of SEQ ID NOS:4-25, the complements of SEQ ID NOS: 4-25, and SEQ ID NOS:26-69.

16. The composition of claim 13, wherein the genotyping oligonucleotide is a primer-extension oligonucleotide.

17. A method for genotyping the immunoglobulin E receptor beta chain (IGERB) gene of an individual, comprising determining for the two copies of the IGERB gene present in the individual the identity of the nucleotide pair at each of[one or more polymorphic sites (PS) selected from] PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.

18. The method of claim 17, wherein the determining step comprises:

(a) isolating from the individual a nucleic acid mixture comprising both copies of the IGERB gene, or a fragment thereof, that are present in the individual;

(b) amplifying from the nucleic acid mixture a target region containing a polymorphic site selected from the group consisting of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13[at least one of the polymorphic sites];

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

19. A method for haplotyping the immunoglobulin E receptor beta chain (IGERB) gene of an

individual which comprises determining, for one copy of the IGERB gene present in the individual, the identity of the nucleotide at each of[one or more polymorphic sites (PS) selected from] PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.

20. The method of claim 19, wherein the determining step comprises

(a) isolating from the individual a nucleic acid molecule containing only one of the two copies of the IGERB gene, or a fragment thereof, that is present in the individual;

(b) amplifying from the nucleic acid molecule a target region containing a polymorphic site selected from the group consisting of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13[at least one of the polymorphic sites];

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

21. A method for predicting a haplotype pair for the immunoglobulin E receptor beta chain (IGERB) gene of an individual comprising:

(a) identifying an IGERB genotype for the individual at each of[two or more of polymorphic sites selected from] PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13;

(b) enumerating all possible haplotype pairs which are consistent with the genotype;

(c) comparing the possible haplotype pairs to the haplotype pair data in Table4[accessing data containing the IGERB haplotype pairs determined in a reference population]; and

(d) assigning a haplotype pair to the individual that is consistent with the data in Table 4.

22. A method for identifying an association between a trait and at least one [genotype or]haplotype or haplotype pair of the immunoglobulin E receptor beta chain gene which comprises comparing the frequency of the [genotype or]haplotype or haplotype pair in a population exhibiting the trait with the frequency of the [genotype or]haplotype or haplotype pair in a reference population, wherein the [genotype or]haplotype or haplotype pair is selected from haplotypes 1-12 in Table 5 and the haplotype pair is selected from the haplotype pairs shown in Table 4[comprises a nucleotide pair or nucleotide located at one or more polymorphic sites selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13], wherein a higher frequency of the [genotype or]haplotype or haplotype pair in the trait population than in the reference population indicates the trait is associated with the [genotype or]haplotype or haplotype pair.

[23 The method of claim 22, wherein the haplotype is selected from haplotype numbers 1-12 shown in Table 5.]

24. The method of claim 22[23], wherein the trait is a clinical response to a drug targeting IGERB .

25. A computer system for storing and analyzing polymorphism data for the immunoglobulin E receptor beta chain gene, comprising:

(a) a central processing unit (CPU);

(b) a communication interface;

(c) a display device;

(d) an input device; and

(e) a database containing the polymorphism data;

wherein the polymorphism data comprises the genotypes and haplotype pairs shown in Table 4 and the haplotypes shown in Table 5.

26. A genome anthology for the immunoglobulin E receptor beta chain (IGERB) gene which comprises IGERB isogenes defined by haplotypes 1-12 shown in Table 5.

27. A method for haplotyping the Immunoglobulin E Receptor I Beta Chain (IGERB) gene of an individual which comprises determining whether the individual has one of the haplotypes represented in Table 5 or one of the haplotype pairs shown in Table 4.

28. The method of claim 27, wherein the determining step comprises identifying the phased sequence of nucleotides present at each of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13 on at least one copy of the individual's IGERB gene.

29. The method of claim 27, wherein the determining step comprises identifying the phased sequence of nucleotides present at each of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13 on both copies of the individual's IGERB gene.

What is Claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) a first nucleotide sequence which is a polymorphic variant of a reference sequence for immunoglobulin E receptor beta chain (IGERB) gene or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:1, and the polymorphic variant comprises an IGERB isogene defined by a haplotype selected from the group consisting of haplotypes 1-12 in Table 5; and

(b) a second nucleotide sequence which is complementary to the first nucleotide sequence.

3. The isolated polynucleotide of claim 1 which is a DNA molecule and comprises both the first and second nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.

4. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 1, wherein the organism expresses an IGERB protein encoded by the first nucleotide sequence.

5. The recombinant organism of claim 4 which is a nonhuman transgenic animal.

6. The isolated polynucleotide of claim 1, wherein the first nucleotide sequence is a polymorphic variant of a fragment of the IGERB isogene, the fragment comprising one or more polymorphisms selected from the group consisting of adenine at PS2, thymine at PS3, cytosine at PS4, cytosine at PS5, thymine at PS6, guanine at PS7, guanine at PS8, guanine at PS9, adenine at PS10, cytosine at PS12 and thymine at PS13.

7. An isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the IGERB cDNA or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:2 and the polymorphic variant comprises a coding sequence of an IGERB isogene defined by a haplotype selected from the group consisting of haplotypes 1-12 in Table 5.

8. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 7, wherein the organism expresses an immunoglobulin E receptor beta chain (IGERB) protein encoded by the polymorphic variant sequence.

9. The recombinant organism of claim 8 which is a nonhuman transgenic animal.

10. An isolated polypeptide comprising an amino acid sequence which is a polymorphic variant of a reference sequence for the IGERB protein or a fragment thereof, wherein the reference sequence comprises SEQ ID NO: 3 and the polymorphic variant is encoded by an isogene defined by one of the haplotypes shown in Table 5.

11. An isolated antibody specific for and immunoreactive with the isolated polypeptide of claim 10.

12. A method for screening for drugs targeting the isolated polypeptide of claim 10 which comprises contacting the IGERB polymorphic variant with a candidate agent and assaying for binding activity.

13. A composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in the immunoglobulin E receptor beta chain (IGERB) gene at a polymorphic site selected from PS2,

PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.

14. The composition of claim 13, wherein the genotyping oligonucleotide is an allele-specific oligonucleotide that specifically hybridizes to an allele of the IGERB gene at a region containing the polymorphic site.

15. The composition of claim 14, wherein the allele-specific oligonucleotide comprises a nucleotide sequence selected from the group consisting of of SEQ ID NOS:4-25, the complements of SEQ ID NOS: 4-25, and SEQ ID NOS:26-69.

16. The composition of claim 13, wherein the genotyping oligonucleotide is a primer-extension oligonucleotide.

17. A method for genotyping the immunoglobulin E receptor beta chain (IGERB) gene of an individual, comprising determining for the two copies of the IGERB gene present in the individual the identity of the nucleotide pair at each of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.

18. The method of claim 17, wherein the determining step comprises:

(a) isolating from the individual a nucleic acid mixture comprising both copies of the IGERB gene, or a fragment thereof, that are present in the individual;

(b) amplifying from the nucleic acid mixture a target region containing a polymorphic site selected from the group consisting of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

19. A method for haplotyping the immunoglobulin E receptor beta chain (IGERB) gene of an individual which comprises determining, for one copy of the IGERB gene present in the individual, the identity of the nucleotide at each of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.

20. The method of claim 19, wherein the determining step comprises

(a) isolating from the individual a nucleic acid molecule containing only one of the two copies of the IGERB gene, or a fragment thereof, that is present in the individual;

(b) amplifying from the nucleic acid molecule a target region containing a polymorphic site selected from the group consisting of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

21. A method for predicting a haplotype pair for the immunoglobulin E receptor beta chain (IGERB) gene of an individual comprising:

(a) identifying an IGERB genotype for the individual at each of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13;

(b) enumerating all possible haplotype pairs which are consistent with the genotype;

(c) comparing the possible haplotype pairs to the haplotype pair data in Table4; and

(d) assigning a haplotype pair to the individual that is consistent with the data in Table 4.

22. A method for identifying an association between a trait and at least one haplotype or haplotype pair of the immunoglobulin E receptor beta chain gene which comprises comparing the frequency of the haplotype or haplotype pair in a population exhibiting the trait with the frequency of the haplotype or haplotype pair in a reference population, wherein the haplotype or haplotype pair is selected from haplotypes 1-12 in Table 5 and the haplotype pair is selected from the haplotype pairs shown in Table 4, wherein a higher frequency of the haplotype or haplotype pair in the trait population than in the reference population indicates the trait is associated with the haplotype or haplotype pair.

24. The method of claim 22, wherein the trait is a clinical response to a drug targeting IGERB .

25. A computer system for storing and analyzing polymorphism data for the immunoglobulin E receptor beta chain gene, comprising:

(a) a central processing unit (CPU);

(b) a communication interface;

(c) a display device;

(d) an input device; and

(e) a database containing the polymorphism data;

wherein the polymorphism data comprises the genotypes and haplotype pairs shown in Table 4 and the haplotypes shown in Table 5.

26. A genome anthology for the immunoglobulin E receptor beta chain (IGERB) gene which comprises IGERB isogenes defined by haplotypes 1-12 shown in Table 5.

27. A method for haplotyping the Immunoglobulin E Receptor I Beta Chain (IGERB) gene of an individual which comprises determining whether the individual has one of the haplotypes represented in Table 5 or one of the haplotype pairs shown in Table 4.

28. The method of claim 27, wherein the determining step comprises identifying the phased sequence

of nucleotides present at each of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13 on at least one copy of the individual's IGERB gene.

29. The method of claim 27, wherein the determining step comprises identifying the phased sequence of nucleotides present at each of PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13 on both copies of the individual's IGERB gene.

DRUG TARGET ISOGENES:
POLYMORPHISMS IN THE IMMUNOGLOBULIN E RECEPTOR BETA CHAIN GENE

RELATED APPLICATIONS

This application claims the benefit of U.S. Provisional Application Serial No. 60/150,423 filed August 24, 1999.

FIELD OF THE INVENTION

This invention relates to variation in genes that encode pharmaceutically important proteins. In particular, this invention provides genetic variants of the human Immunoglobulin E receptor beta chain (IGERB) gene and methods for identifying which variant(s) of this gene is/are possessed by an individual.

BACKGROUND OF THE INVENTION

Current methods for identifying pharmaceuticals to treat disease often start by identifying, cloning, and expressing an important target protein related to the disease. A determination of whether an agonist or antagonist is needed to produce an effect that may benefit a patient with the disease is then made. Then, vast numbers of compounds are screened against the target protein to find new potential drugs. The desired outcome of this process is a drug that is specific for the target, thereby reducing the incidence of the undesired side effects usually caused by a compound's activity at non-intended targets.

What this approach fails to consider, however, is that natural variability exists in any and every population with respect to a particular protein. A target protein currently used to screen drugs typically is expressed by a gene cloned from an individual who was arbitrarily selected. However, the nucleotide sequence of a particular gene may vary tremendously among individuals. Subtle alteration(s) in the primary nucleotide sequence of a gene encoding a target protein may be manifested as significant variation in expression of or in the structure and/or function of the protein. Such alterations may explain the relatively high degree of uncertainty inherent in treatment of individuals with drugs whose design is based upon a single representative example of the target. For example, it is well-established that some classes of drugs frequently have lower efficacy in some individuals than others, which means such individuals and their physicians must weigh the possible benefit of a larger dosage against a greater risk of side effects. In addition, variable information on the biological function or effects of a particular protein may be due to different scientists unknowingly studying different isoforms of the gene encoding the protein. Thus, information on the type and frequency of genomic variation that exists for pharmaceutically important proteins would be useful.

The organization of single nucleotide variations (polymorphisms) in the primary sequence of a gene into one of the limited number of combinations that exist as units of inheritance is termed a haplotype. Each haplotype therefore contains significantly more information than individual

haplotype. Each haplotype therefore contains significantly more information than individual unorganized polymorphisms. Haplotypes provide an accurate measurement of the genomic variation in the two chromosomes of an individual.

It is well-established that many diseases are associated with specific variations in gene sequences. However while there are examples in which individual polymorphisms act as genetic markers for a particular phenotype, in other cases an individual polymorphism may be found in a variety of genomic backgrounds and therefore shows no definitive coupling between the polymorphism and the causative site for the phenotype (Clark AG et al. 1998 *Am J Hum Genet* **63**:595-612; Ulbrecht M et al. 2000 *Am J Respir Crit Care Med* **161**: 469-74). In addition, the marker may be predictive in some populations, but not in other populations (Clark AG et al. 1998 *supra*). In these instances, a haplotype will provide a superior genetic marker for the phenotype (Clark AG et al. 1998 *supra*; Ulbrecht M et al. 2000, *supra*; Ruaño G & Stephens JC *Gen Eng News* **19** (21), December 1999).

Analysis of the association between each observed haplotype and a particular phenotype permits ranking of each haplotype by its statistical power of prediction for the phenotype. Haplotypes found to be strongly associated with the phenotype can then have that positive association confirmed by alternative methods to minimize false positives. For a gene suspected to be associated with a particular phenotype, if no observed haplotypes for that gene show association with the phenotype of interest, then it may be inferred that variation in the gene has little, if any, involvement with that phenotype (Ruaño & Stephens 1999, *supra*). Thus, information on the observed haplotypes and their frequency of occurrence in various population groups will be useful in a variety of research and clinical applications.

One possible drug target for the treatment of diseases involving immune response is the Immunoglobulin E receptor beta chain (IGERB) gene or its encoded product. The high affinity IgE receptor (IGER) belongs to the family of antibody Fc receptors that play an important role in the immune response by coupling the specificity of secreted antibodies to a variety of cells of the immune system (OMIM: 147138). Fc receptors initiate immune system reactions in normal immunity, allergies, antibody-mediated tumor recognition, and autoimmune diseases such as arthritis. IGER mediates IgE-dependent peripheral and systemic anaphylaxis, regulates IgE metabolism, and plays a role in the growth and differentiation of various cells of the immune system.

IGER initiates the immediate hypersensitivity response from mast cells and basophils, and evidence indicates that this receptor is involved in antiparasitic reactions from platelet and eosinophils, and in antigen delivery to dendritic cells for MHC class II presentation pathways activating T cells. Moreover, IGER exerts a regulatory effect on IgE production, as well as differentiation and growth of mast cell and B-lymphocytes. Stimulation of IGER initiates a cascade of events resulting in a number of cellular events. For example, mast cells release inflammatory mediators, such as histamine. Cytokines are also released, particularly interleukin 4 (IL-4), which is critical in the B-cell switching and IgE synthesis pathways, as well as a feed-back up-regulation of IGER synthesis. IGER stimulation also induces expression and functions of other mast cell surface receptors, such as CD40, involved in

immune cell growth and differentiation, as well as IgE metabolism. Other factors whose expression and/or secretion are regulated by IGER include, interleukin 6 (IL-6), tissue necrosis factor alpha (TNFα), RANTES, and serotonin, among others.

IGER is a tetrameric transmembrane protein containing an alpha, beta, and two disulfide-bonded gamma polypeptides (gamma subunit). The alpha subunit, IGERA, binds IgE with high affinity (Kd ~109-1010M) and can be secreted as a soluble IgE-binding fragment. The gamma subunit, IGERG, mediates receptor assembly and signal transduction, and is a common component of other Fc receptors, including the high-affinity and low-affinity IgG receptors, and the TCR/CD3 T-cell receptor complex. The role of the beta subunit, IGERB, is more enigmatic, although it is also involved in signal transduction and receptor autophosphorylation. IGERB, also known as FCER1B, is essential for full activation of mast cells for the allergic response and is an amplifier of signaling from the gamma subunit.

The alpha subunit of IGER consists of a C-terminal cytoplasmic tail, a single transmembrane region, and an N-terminal extracellular region divided into two large immunoglobulin (Ig) domains. The Ig domains are each 85 amino acids in length, and are bent at an acute angle to form a convex binding site for IgE. The second domain has a prominent loop that projects above the domain and is a site of interaction with IgE. IGERB is a four transmembrane protein with N-terminal and C-terminal cytoplasmic tails. The N-terminal cytoplasmic domain interacts with the cytoplasmic domains of the IGERG subunits. The C-terminal cytoplasmic tail of IGERB associates with the cytoplasmic tail of the alpha subunit. IGERG has a short extracellular N-terminal tail, a single transmembrane region, and a C-terminal cytoplasmic domain.

Both IGERB and IGERG have immunoreceptor tyrosine activation motifs (ITAM) in their cytoplasmic domains. The IGERB ITAM appears in the C-terminal cytoplasmic domain. Evidence suggests that the two ITAM domains act synergistically, associating with specific protein tyrosine kinases that are capable of triggering cell activation via protein-tyrosine phosphorylation. Receptor subunit cross-linking activates the src kinase, Lyn, associated with the IGERB ITAM, in turn phosphorylating two tyrosine residues in the ITAM. This event activates the src kinase, Syk, associated with the IGERG ITAM, phosphorylating the ITAM tyrosines in that subunit. Deletion of the C-terminal cytoplasmic domain of IGERB, containing the Lyn ITAM, results in an inactive receptor. Mutation of either or both tryosines in the IGERB ITAM results in non-phosphorylation of IGERB and IGERG tyrosines.

The gene for the beta subunit of the IgE receptor is located on human chromosome 11q13 (Szepetowski and Gaudray, *Genomics* 19:399-400, 1989; Young et al., *J. Med. Genet.* 29:236-238 1992; Sandford et al., *Lancet* 341:332-334, 1993). The gene spans approximately 10 kilobase pairs (kb) of genomic DNA and consists of seven exons encoding 244 amino acids (Kuster et al., *J. Biol. Chem.* 267:12782-12787, 1992). Reference sequences for the IGERB gene, comprising 11,298 nucleotides (SEQ ID NO:1; GenBank Accession No. M89796; Kuster et al., *supra*), coding sequence, and protein

(SEQ ID NO:3; GenBank Accession No. AAA60269; Kuster et al., *supra*), are shown in Figs. 1, 2, and 3, respectively. Significant features reported for the IGERB gene and its encoded protein include: a canonical TATA box located at nucleotide position 412, a 5′ untranslated sequence comprising the first 102 nucleotides of exon 1; an initiation codon at nucleotide position 456; an N-terminal cytoplasmic tail of the protein encoded by the remainder of exon 1 and a portion of exon 2; four transmembrane (TM) regions encoded by exons 2 and 3 (TM-1), exons 3 and 4 (TM-2), exon 5 (TM-3), and exon 6 (TM-4); a C-terminal cytoplasmic tail and 3′ untranslated region encoded by exon 7; and an ITAM motif in the C-terminal cytoplasmic domain (amino acids 217-232; Fig. 3).

The primary interest in the IGERB gene derives from linkage and association studies of the genetic components of atopy. Atopy is a common familial disorder caused by genetic and environmental factors. Atopy is characterized by exaggerated T- helper cell type II lymphocyte responses to common allergens, such as pollens and dust mites, with sustained, enhanced production of IgE. Allergy, asthma, rhinitis, and eczema are atopic hypersensitivity diseases. IgE binds to the high affinity IgE receptor presented on mucosal mast cells and basophils. IgE binding of allergens activates the receptor and initiates a cascade, leading to cellular release of inflammatory mediators. Dysregulation of the normal immediate hypersensitivity response results in abnormally high and sustained IgE serum levels and leads to mucosal inflammation. Atopy is detected by elevated total serum IgE levels, positive skin prick tests to common allergens, and specific serum IgE against these allergens. All three have been strongly correlated with each other and the presence of the symptoms of allergic reaction-- wheezing, coughing, sneezing, and nasal blockage. Approximately 20% of the world population is affected by allergies, with over 50% of western populations testing positive to skin prick tests of one or more common allergens. Up to 10% of children suffer from atopic asthma, accounting for approximately one-third of US pediatric emergency room visits. While a single genetic determinant is unlikely to be the causative factor in asthma, allergy, or other atopic diseases, therapeutics aimed at the obligatory binding of IgE to IGER for initiation of allergic response could provide a single treatment for the various manifestations of atopic hypersensitivity.

Recent studies have suggested that polymorphisms in the IGERB gene are associated with genetic predisposition to atopy and/or elevated serum IgE. Linkage between IgE responsiveness and other atopic indicators was linked to a locus on chromosome 11q (Cookson et al., *Lancet* 1: 1292-1295, 1989). Sandford et al., *supra*, used a CA microsatellite repeat (STR) located in the fifth intron (Fig 1; nucleotides 5483-5512) to map the IGERB gene to 11q13, and also found linkage between this gene and clinical atopy. Shirakawa et al (*Lancet* 347: 394-395, 1996) identified an RsaI restriction fragment length polymorphism (RFLP) in the second intron, as well as three single nucleotide polymorphisms (SNPs) in the sixth exon of the gene (Shirakawa et al., *Nature Genet.* 7:125-130; Shirakawa et al., *Hum. Mol. Genet.* 5:1129-1130). The exact site of the RsaI RFLP was not published, but association with atopic asthma, rhinitis, and eczema was found in a Japanese clinical population. The three SNPs, at nucleotide positions 5643 (A or T), 5645 (T or G), and 5649 (G or T) (Fig 1), result in amino acid

variation in the encoded protein that are reportedly strongly associated with maternal transmission of atopy. Specifically, when nucleotides 5643 and 5645 are adenine and thymine, respectively, position 181 of the encoded IGERB is isoleucine (Fig.3), whereas when these positions are thymine and guanine, respectively, the IGERB has leucine at amino acid 181. The third SNP at nucleotide position 5649 results in variation between leucine and valine at amino acid position 183. Interestingly, this V183L variant has never been observed by itself, but is always accompanied by the I181L variant. In contrast, the I181L variant alone was found frequently. These amino acid residues at 181 and 183 form part of the fourth transmembrane domain. Mutagenicity studies on IGERB and other multiple transmembrane domain proteins have shown that amino acid substitutions in these regions can have significant effects on protein expression and function.

Hill and Cookson (*Hum. Mol. Genet.* 5: 959-962, 1996) identified a SNP of adenine or guanine at nucleotide position 7297 in the seventh exon (Fig. 1) which leads to variation between glutamic acid and glycine at amino acid 237 (Fig. 3) located in the C-terminal cytoplasmic tail close to the ITAM domain. This group found a significant association between the Gly237 allele and susceptibility to asthma. The Glu237Gly substitution is predicted to introduce a change in hydrophobicity of that region of the receptor, and thus may have functional consequences due to the amino acid's proximity to the ITAM domain, which is necessary for receptor phosphorylation.

A recent report disclosed the identification of two silent SNPs, one in the promoter region (cytosine or thymine, nucleotide position 245) and in the third intron (thymine or guanine, nucleotide position 4392) (Dickson et al., *Thorax* 54:409-412, 1999). This report found no association between these two SNPs, alone or together, and asthma, atopy, or bronchial hypersensitivity. Recent studies suggest that the cytosine or thymine polymorphism corresponding to nucleotide position 245 is associated with elevated total serum IgE levels in Japanese patients with asthma (Hizawa et al. Am J Respir Crit Care Med. 2000 Mar;161(3 Pt 1):906-9).

Palmer et al (*Am. J. Hum. Genet.* 61:182-188, 1997) described a polymorphic site in the 3′ untranslated region of exon 7 (nucleotide position 9867). This polymorphism is associated with total serum IgE levels in endemically parasitzed Australian Aborigines. In the same study, a similar association was found with the intron 5 STR, indicating that variation at the IGERB locus may regulate IgE mediated immune response to parasitic infection.

Because of the potential for polymorphisms in the IGERB gene to affect the expression and function of the encoded protein, it would be useful to determine whether additional polymorphisms exist in the IGERB gene, as well as how such polymorphisms are combined in different copies of the gene. Such information would be useful for studying the biological function of IGERB as well as in identifying drugs targeting this protein for the treatment of disorders related to its abnormal expression or function.

SUMMARY OF THE INVENTION

Accordingly, the inventors herein have discovered 11 novel polymorphic sites in the IGERB gene. These polymorphic sites (PS) correspond to the following nucleotide positions in the indicated GenBank Accession Number: 266 (PS2), 1456 (PS3), 2253 (PS4), 2302 (PS5), 5128 (PS6), 5173 (PS7), 5232 (PS8), 5252 (PS9), 5256 (PS10), 7402 (PS12) and 7446 (PS13) in M89796.1. The polymorphisms at these sites are guanine or adenine at PS2, cytosine or thymine at PS3, guanine or cytosine at PS4, thymine or cytosine at PS5, cytosine or thymine at PS6, adenine or guanine at PS7, thymine or guanine at PS8, thymine or guanine at PS9, guanine or adenine at PS10, thymine or cytosine at PS12 and cytosine or thymine at PS13. In addition, the inventors have determined the identity of the alternative nucleotides present at these sites, as well as at the previously identified sites at nucleotides 245 (PS1) and 7297 (PS11) in M89796.1. It is believed that IGERB-encoding polynucleotides containing one or more of the novel polymorphic sites reported herein will be useful in studying the expression and biological function of IGERB, as well as in developing drugs targeting this protein. In addition, information on the combinations of polymorphisms in the IGERB gene may have diagnostic and forensic applications.

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the IGERB gene or a fragment thereof. The reference sequence comprises SEQ ID NO:1 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of adenine at PS2, thymine at PS3, cytosine at PS4, cytosine at PS5, thymine at PS6, guanine at PS7, guanine at PS8, guanine at PS9, adenine at PS10, cytosine at PS12 and thymine at PS13. In a preferred embodiment, the polymorphic variant comprises one or more additional polymorphisms selected from the group consisting of thymine at PS1 and guanine at PS11. A particularly preferred polymorphic variant is a naturally-occurring isoform (also referred to herein as an "isogene") of the IGERB gene. An IGERB isogene of the invention comprises cytosine or thymine at PS1, guanine or adenine at PS2, cytosine or thymine at PS3, guanine or cytosine at PS4, thymine or cytosine at PS5, cytosine or thymine at PS6, adenine or guanine at PS7, thymine or guanine at PS8, thymine or guanine at PS9, guanine or adenine at PS10, adenine or guanine at PS11, thymine or cytosine at PS12 and cytosine or thymine at PS13. The invention also provides a collection of IGERB isogenes, referred to herein as an IGERB genome anthology.

An IGERB isogene may be defined by the combination and order of these polymorphisms in the isogene, which is referred to herein as an IGERB haplotype. Thus, the invention also provides data on the number of different IGERB haplotypes found in the above four population groups. This haplotype data is useful in methods for deriving an IGERB haplotype from an individual's genotype for the IGERB gene and for determining an association between an IGERB haplotype and a particular trait.

In another embodiment, the invention provides a polynucleotide comprising a polymorphic variant of a reference sequence for an IGERB cDNA or a fragment thereof. The reference sequence comprises SEQ ID NO:2 (Fig. 2) and the polymorphic cDNA comprises at least one polymorphism

selected from the group consisting of thymine at a position corresponding to nucleotide 132, thymine at a position corresponding to nucleotide 428, guanine at a position corresponding to nucleotide 473, and guanine at a position corresponding to nucleotide 532 . In a preferred embodiment, the polymorphic variant comprises an additional polymorphism of guanine at a position corresponding to nucleotide 710.

Polynucleotides complementary to these IGERB genomic and cDNA variants are also provided by the invention.

In other embodiments, the invention provides a recombinant expression vector comprising one of the polymorphic genomic variants operably linked to expression regulatory elements as well as a recombinant host cell transformed or transfected with the expression vector. The recombinant vector and host cell may be used to express IGERB for protein structure analysis and drug binding studies.

In yet another embodiment, the invention provides a polypeptide comprising a polymorphic variant of a reference amino acid sequence for the IGERB protein. The reference amino acid sequence comprises SEQ ID NO:3 (Fig. 3) and the polymorphic variant comprises at least one variant amino acid selected from the group consisting of methionine at a position corresponding to amino acid position 143, cysteine at a position corresponding to amino acid position 158, and alanine at a position corresponding to amino acid position 178 . In some embodiments, the polymorphic variant also comprises glycine at a position corresponding to amino acid position 237. A polymorphic variant of IGERB is useful in studying the effect of the variation on the biological activity of IGERB as well as studying the binding affinity of candidate drugs targeting IGERB for the treatment of diseases involving immune response.

The present invention also provides antibodies that recognize and bind to the above polymorphic IGERB protein variant. Such antibodies can be utilized in a variety of diagnostic and prognostic formats and therapeutic methods.

In other embodiments, the invention provides methods, compositions, and kits for haplotyping and/or genotyping the IGERB gene in an individual. The methods involve identifying the nucleotide or nucleotide pair present at one or more polymorphic sites selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13 in one or both copies of the IGERB gene from the individual. The compositions contain oligonucleotide probes and primers designed to specifically hybridize to one or more target regions containing, or that are adjacent to, a polymorphic site. The methods and compositions for establishing the genotype or haplotype of an individual at the novel polymorphic sites described herein are useful for studying the effect of the polymorphisms in the etiology of diseases affected by the expression and function of the IGERB protein, studying the efficacy of drugs targeting IGERB, predicting individual susceptibility to diseases affected by the expression and function of the IGERB protein and predicting individual responsiveness to drugs targeting IGERB.

In yet another embodiment, the invention provides a method for identifying an association between a genotype or haplotype and a trait. In preferred embodiments, the trait is susceptibility to a disease, severity of a disease, the staging of a disease or response to a drug. Such methods have

applicability in developing diagnostic tests and therapeutic treatments for diseases involving immune response.

The present invention also provides transgenic animals comprising one of the IGERB genomic polymorphic variants described herein and methods for producing such animals. The transgenic animals are useful for studying expression of the IGERB isogenes *in vivo*, for *in vivo* screening and testing of drugs targeted against IGERB protein, and for testing the efficacy of therapeutic agents and compounds for diseases involving immune response in a biological system.

The present invention also provides a computer system for storing and displaying polymorphism data determined for the IGERB gene. The computer system comprises a computer processing unit; a display; and a database containing the polymorphism data. The polymorphism data includes the polymorphisms, the genotypes and the haplotypes identified for the IGERB gene in a reference population. In a preferred embodiment, the computer system is capable of producing a display showing IGERB haplotypes organized according to their evolutionary relationships.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 illustrates a reference sequence for the IGERB gene (Genbank Version Number M89796.1.1; contiguous lines; SEQ ID NO:1), with the start and stop positions of each region of coding sequence indicated below the sequence by the numbers within the brackets and the polymorphic sites and polymorphisms identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence.

Figure 2 illustrates a reference sequence for the IGERB coding sequence (contiguous lines; SEQ ID NO:2), with the polymorphic sites and polymorphisms identified by Applicants in a reference population indicated by the variant nucleotide positioned below the polymorphic site in the sequence.

Figure 3illustrates a reference sequence for the IGERB protein (contiguous lines; SEQ ID NO:3), with the variant amino acids caused by the polymorphisms of Fig. 2 positioned below the polymorphic site in the sequence.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

The present invention is based on the discovery of novel variants of the IGERB gene. As described in more detail below, the inventors herein discovered 11 novel polymorphic sites by characterizing the IGERB gene found in genomic DNAs isolated from Index Repository IA that contains immortalized cell lines from one chimpanzee and 93 human individuals and Index Repository IB that contains 70 human individuals. These two repositories contain 10 individuals in common.

The human individuals in Index Repository IA included a reference population of 79 unrelated individuals self-identified as belonging to one of four major population groups: Caucasian (22 individuals), African descent (20 individuals) Asian (20 individuals) Hispanic/Latino (17 individuals). To the extent possible, the members of this reference population were organized into population

subgroups by the self-identified ethnogeographic origin of their four grandparents as shown in Table 1 below. In addition, Index Repository IA contains three unrelated indigenous American Indians (one from each of North, Central and South America), one three-generation Caucasian family (from the CEPH Utah cohort) and one two-generation African-American family.

Table 1. Population Groups in the Index Repository 1A

Population Group	Population Subgroup	No. of Individuals
African descent		20
	Sierra Leone	1
Asian		20
	Burma	1
	China	3
	Japan	6
	Korea	1
	Philippines	5
	Vietnam	4
Caucasian		22
	British Isles	3
	British Isles/Central	4
	British Isles/Eastern	1
	Central/Eastern	1
	Eastern	3
	Central/Mediterranean	1
	Mediterranean	2
	Scandinavian	2
Hispanic/Latino		17
	Caribbean	7
	Caribbean (Spanish Descent)	2
	Central American (Spanish Descent)	1
	Mexican American	4
	South American (Spanish Descent)	3

Index Repository IB contains a reference population of 70 human individuals comprised of 4 three-generation families (from the CEPH Utah cohort) as well as unrelated African-American, Asian and Caucasian individuals. A total of 38 individuals in this reference population are unrelated.

Using the IGERB genotypes identified in Index Repository 1A and the methodology described in the Examples below, the inventors herein also determined the haplotypes found on each chromosome for most human members of this repository. The IGERB genotypes and haplotypes found in Index Repository 1A include those shown in Tables 4 and 5, respectively. The polymorphism and haplotype data disclosed herein are useful for studying population diversity, anthropological lineage, the significance of diversity and lineage at the phenotypic level, paternity testing, forensic applications, and for identifying associations between the IGERB genetic variation and a trait such as level of drug response or susceptibility to disease.

In the context of this disclosure, the following terms shall be defined as follows unless otherwise indicated:

Allele - A particular form of a genetic locus, distinguished from other forms by its particular nucleotide sequence.

Candidate Gene – A gene which is hypothesized to be responsible for a disease, condition, or the response to a treatment, or to be correlated with one of these.

Gene - A segment of DNA that contains all the information for the regulated biosynthesis of an RNA product, including promoters, exons, introns, and other untranslated regions that control expression.

Genotype – An unphased 5′ to 3′ sequence of nucleotide pair(s) found at one or more polymorphic sites in a locus on a pair of homologous chromosomes in an individual. As used herein, genotype includes a full-genotype and/or a sub-genotype as described below.

Full-genotype – The unphased 5′ to 3′ sequence of nucleotide pairs found at all known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Sub-genotype – The unphased 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a pair of homologous chromosomes in a single individual.

Genotyping – A process for determining a genotype of an individual.

Haplotype – A 5′ to 3′ sequence of nucleotides found at one or more polymorphic sites in a locus on a single chromosome from a single individual. As used herein, haplotype includes a full-haplotype and/or a sub-haplotype as described below.

Full-haplotype – The 5′ to 3′ sequence of nucleotides found at all known polymorphic sites in a locus on a single chromosome from a single individual.

Sub-haplotype – The 5′ to 3′ sequence of nucleotides seen at a subset of the known polymorphic sites in a locus on a single chromosome from a single individual.

Haplotype pair – The two haplotypes found for a locus in a single individual.

Haplotyping – A process for determining one or more haplotypes in an individual and includes use of family pedigrees, molecular techniques and/or statistical inference.

Haplotype data - Information concerning one or more of the following for a specific gene: a listing of the haplotype pairs in each individual in a population; a listing of the different haplotypes in a population; frequency of each haplotype in that or other populations, and any known associations between one or more haplotypes and a trait.

Isoform – A particular form of a gene, mRNA, cDNA or the protein encoded thereby, distinguished from other forms by its particular sequence and/or structure.

Isogene – One of the isoforms of a gene found in a population. An isogene contains all of the polymorphisms present in the particular isoform of the gene.

Isolated – As applied to a biological molecule such as RNA, DNA, oligonucleotide, or protein, isolated means the molecule is substantially free of other biological molecules such as nucleic acids, proteins, lipids, carbohydrates, or other material such as cellular debris and growth media. Generally, the term "isolated" is not intended to refer to a complete absence of such material or to absence of water, buffers, or salts, unless they are present in amounts that substantially interfere with the methods of the present invention.

Locus - A location on a chromosome or DNA molecule corresponding to a gene or a physical or phenotypic feature.

Naturally-occurring – A term used to designate that the object it is applied to, e.g., naturally-occurring polynucleotide or polypeptide, can be isolated from a source in nature and which has not been intentionally modified by man.

Nucleotide pair – The nucleotides found at a polymorphic site on the two copies of a chromosome from an individual.

Phased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, phased means the combination of nucleotides present at those polymorphic sites on a single copy of the locus is known.

Polymorphic site (PS) – A position within a locus at which at least two alternative sequences are found in a population, the most frequent of which has a frequency of no more than 99%.

Polymorphic variant – A gene, mRNA, cDNA, polypeptide or peptide whose nucleotide or amino acid sequence varies from a reference sequence due to the presence of a polymorphism in the gene.

Polymorphism – The sequence variation observed in an individual at a polymorphic site. Polymorphisms include nucleotide substitutions, insertions, deletions and microsatellites and may, but need not, result in detectable differences in gene expression or protein function.

Polymorphism data – Information concerning one or more of the following for a specific gene: location of polymorphic sites; sequence variation at those sites; frequency of polymorphisms in one or more populations; the different genotypes and/or haplotypes determined for the gene; frequency of one or more of these genotypes and/or haplotypes in one or more populations; any known association(s) between a trait and a genotype or a haplotype for the gene.

Polymorphism Database – A collection of polymorphism data arranged in a systematic or methodical way and capable of being individually accessed by electronic or other means.

Polynucleotide – A nucleic acid molecule comprised of single-stranded RNA or DNA or comprised of complementary, double-stranded DNA.

Population Group – A group of individuals sharing a common ethnogeographic origin.

Reference Population – A group of subjects or individuals who are predicted to be representative of the genetic variation found in the general population. Typically, the reference population represents the genetic variation in the population at a certainty level of at least 85%, preferably at least 90%, more preferably at least 95% and even more preferably at least 99%.

Single Nucleotide Polymorphism (SNP) – Typically, the specific pair of nucleotides observed at a single polymorphic site. In rare cases, three or four nucleotides may be found.

Subject – A human individual whose genotypes or haplotypes or response to treatment or disease state are to be determined.

Treatment - A stimulus administered internally or externally to a subject.

Unphased – As applied to a sequence of nucleotide pairs for two or more polymorphic sites in a locus, unphased means the combination of nucleotides present at those polymorphic sites on a single

copy of the locus is not known.

The inventors herein have discovered 11 novel polymorphic sites in the IGERB gene. The polymorphic sites identified by the inventors are referred to as PS1-13 to designate the order in which they are located in the gene (see Table 3 below), with the novel polymorphic sites referred to as PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.

Thus, in one embodiment, the invention provides an isolated polynucleotide comprising a polymorphic variant of the IGERB gene or a fragment of the gene which contains at least one of the novel polymorphic sites described herein. The nucleotide sequence of a variant IGERB gene is identical to the reference genomic sequence for those portions of the gene examined, as described in the Examples below, except that it comprises a different nucleotide at one or more of the novel polymorphic sites PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13, and may also comprise one or more additional polymorphisms selected from the group consisting of thymine at PS1 and guanine at PS11 . Similarly, the nucleotide sequence of a variant fragment of the IGERB gene is identical to the corresponding portion of the reference sequence except for having a different nucleotide at one or more of the novel polymorphic sites described herein. Thus, the invention specifically does not include polynucleotides comprising a nucleotide sequence identical to the reference sequence (or other reported IGERB sequences) or to portions of the reference sequence (or other reported IGERB sequences), except for genotyping oligonucleotides as described below.

The location of a polymorphism in a variant gene or fragment is identified by aligning its sequence against SEQ ID NO:1. The polymorphism is selected from the group consisting of adenine at PS2, thymine at PS3, cytosine at PS4, cytosine at PS5, thymine at PS6, guanine at PS7, guanine at PS8, guanine at PS9, adenine at PS10, cytosine at PS12 and thymine at PS13. In a preferred embodiment, the polymorphic variant comprises a naturally-occurring isogene of the IGERB gene which is defined by any one of haplotypes 1-12 shown in Table 5 below.

Polymorphic variants of the invention may be prepared by isolating a clone containing the IGERB gene from a human genomic library. The clone may be sequenced to determine the identity of the nucleotides at the polymorphic sites described herein. Any particular variant claimed herein could be prepared from this clone by performing *in vitro* mutagenesis using procedures well-known in the art.

IGERB isogenes may be isolated using any method that allows separation of the two "copies" of the IGERB gene present in an individual, which, as readily understood by the skilled artisan, may be the same allele or different alleles. Separation methods include targeted *in vivo* cloning (TIVC) in yeast as described in WO 98/01573, U.S. Patent No. 5,866,404, and U.S. Patent No. 5,972,614. Another method, which is described in U.S. Patent No. 5,972,614, uses an allele specific oligonucleotide in combination with primer extension and exonuclease degradation to generate hemizygous DNA targets. Yet other methods are single molecule dilution (SMD) as described in Ruaño et al., Proc. Natl. Acad. Sci. 87:6296-6300, 1990; and allele specific PCR (Ruaño et al., 17 Nucleic Acids. Res. 8392, 1989; Ruaño et al., 19 Nucleic Acids Res. 6877-6882, 1991; Michalatos-Beloin et al., 24 Nucleic Acids Res.

4841-4843, 1996).

The invention also provides IGERB genome anthologies, which are collections of IGERB isogenes found in a given population. The population may be any group of at least two individuals, including but not limited to a reference population, a population group, a family population, a clinical population, and a same sex population. An IGERB genome anthology may comprise individual IGERB isogenes stored in separate containers such as microtest tubes, separate wells of a microtitre plate and the like. Alternatively, two or more groups of the IGERB isogenes in the anthology may be stored in separate containers. Individual isogenes or groups of isogenes in a genome anthology may be stored in any convenient and stable form, including but not limited to in buffered solutions, as DNA precipitates, freeze-dried preparations and the like. A preferred IGERB genome anthology of the invention comprises a set of isogenes defined by the haplotypes shown in Table 5 below.

An isolated polynucleotide containing a polymorphic variant nucleotide sequence of the invention may be operably linked to one or more expression regulatory elements in a recombinant expression vector capable of being propagated and expressing the encoded IGERB protein in a prokaryotic or a eukaryotic host cell. Examples of expression regulatory elements which may be used include, but are not limited to, the lac system, operator and promoter regions of phage lambda, yeast promoters, and promoters derived from vaccinia virus, adenovirus, retroviruses, or SV40. Other regulatory elements include, but are not limited to, appropriate leader sequences, termination codons, polyadenylation signals, and other sequences required for the appropriate transcription and subsequent translation of the nucleic acid sequence in a given host cell. Of course, the correct combinations of expression regulatory elements will depend on the host system used. In addition, it is understood that the expression vector contains any additional elements necessary for its transfer to and subsequent replication in the host cell. Examples of such elements include, but are not limited to, origins of replication and selectable markers. Such expression vectors are commercially available or are readily constructed using methods known to those in the art (e.g., F. Ausubel et al., 1987, in "Current Protocols in Molecular Biology", John Wiley and Sons, New York, New York). Host cells which may be used to express the variant IGERB sequences of the invention include, but are not limited to, eukaryotic and mammalian cells, such as animal, plant, insect and yeast cells, and prokaryotic cells, such as E. coli, or algal cells as known in the art. The recombinant expression vector may be introduced into the host cell using any method known to those in the art including, but not limited to, microinjection, electroporation, particle bombardment, transduction, and transfection using DEAE-dextran, lipofection, or calcium phosphate (see e.g., Sambrook et al. (1989) in "Molecular Cloning. A Laboratory Manual", Cold Spring Harbor Press, Plainview, New York). In a preferred aspect, eukaryotic expression vectors that function in eukaryotic cells, and preferably mammalian cells, are used. Non-limiting examples of such vectors include vaccinia virus vectors, adenovirus vectors, herpes virus vectors, and baculovirus transfer vectors. Preferred eukaryotic cell lines include COS cells, CHO cells, HeLa cells, NIH/3T3 cells, and embryonic stem cells (Thomson, J. A. et al., 1998 Science 282:1145-1147). Particularly

preferred host cells are mammalian cells.

As will be readily recognized by the skilled artisan, expression of polymorphic variants of the IGERB gene will produce IGERB mRNAs varying from each other at any polymorphic site retained in the spliced and processed mRNA molecules. These mRNAs can be used for the preparation of an IGERB cDNA comprising a nucleotide sequence which is a polymorphic variant of the IGERB reference coding sequence shown in Figure 2. Thus, the invention also provides IGERB mRNAs and corresponding cDNAs which comprise a nucleotide sequence that is identical to SEQ ID NO:2 (Fig. 2), or its corresponding RNA sequence, except for having one or more polymorphisms selected from the group consisting of thymine at a position corresponding to nucleotide 132, thymine at a position corresponding to nucleotide 428, guanine at a position corresponding to nucleotide 473, and guanine at a position corresponding to nucleotide 532, and may also comprise an additional polymorphism of guanine at a position corresponding to nucleotide 710. Fragments of these variant mRNAs and cDNAs are included in the scope of the invention, provided they contain the novel polymorphisms described herein. The invention specifically excludes polynucleotides identical to previously identified and characterized IGERB cDNAs and fragments thereof. Polynucleotides comprising a variant RNA or DNA sequence may be isolated from a biological sample using well-known molecular biological procedures or may be chemically synthesized.

Genomic and cDNA fragments of the invention comprise at least one novel polymorphic site identified herein and have a length of at least 10 nucleotides and may range up to the full length of the gene. Preferably, a fragment according to the present invention is between 100 and 3000 nucleotides in length, and more preferably between 200 and 2000 nucleotides in length, and most preferably between 500 and 1000 nucleotides in length.

In describing the polymorphic sites identified herein, reference is made to the sense strand of the gene for convenience. However, as recognized by the skilled artisan, nucleic acid molecules containing the IGERB gene may be complementary double stranded molecules and thus reference to a particular site on the sense strand refers as well to the corresponding site on the complementary antisense strand. Thus, reference may be made to the same polymorphic site on either strand and an oligonucleotide may be designed to hybridize specifically to either strand at a target region containing the polymorphic site. Thus, the invention also includes single-stranded polynucleotides which are complementary to the sense strand of the IGERB genomic variants described herein.

Polynucleotides comprising a polymorphic gene variant or fragment may be useful for therapeutic purposes. For example, where a patient could benefit from expression, or increased expression, of a particular IGERB protein isoform, an expression vector encoding the isoform may be administered to the patient. The patient may be one who lacks the IGERB isogene encoding that isoform or may already have at least one copy of that isogene.

In other situations, it may be desirable to decrease or block expression of a particular IGERB isogene. Expression of an IGERB isogene may be turned off by transforming a targeted organ, tissue or

cell population with an expression vector that expresses high levels of untranslatable mRNA for the isogene. Alternatively, oligonucleotides directed against the regulatory regions (e.g., promoter, introns, enhancers, 3′ untranslated region) of the isogene may block transcription. Oligonucleotides targeting the transcription initiation site, e.g., between positions −10 and +10 from the start site are preferred. Similarly, inhibition of transcription can be achieved using oligonucleotides that base-pair with region(s) of the isogene DNA to form triplex DNA (see e.g., Gee et al. in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing Co., Mt. Kisco, N.Y., 1994). Antisense oligonucleotides may also be designed to block translation of IGERB mRNA transcribed from a particular isogene. It is also contemplated that ribozymes may be designed that can catalyze the specific cleavage of IGERB mRNA transcribed from a particular isogene.

The oligonucleotides may be delivered to a target cell or tissue by expression from a vector introduced into the cell or tissue *in vivo* or *ex vivo*. Alternatively, the oligonucleotides may be formulated as a pharmaceutical composition for administration to the patient. Oligoribonucleotides and/or oligodeoxynucleotides intended for use as antisense oligonucleotides may be modified to increase stability and half-life. Possible modifications include, but are not limited to phosphorothioate or 2′ O-methyl linkages, and the inclusion of nontraditional bases such as inosine and queosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytosine, guanine, thymine, and uracil which are not as easily recognized by endogenous nucleases.

The invention also provides an isolated polypeptide comprising a polymorphic variant of the reference IGERB amino acid sequence shown in Figure 3. The location of a variant amino acid in an IGERB polypeptide or fragment of the invention is identified by aligning its sequence against SEQ ID NO:3 (Fig. 3). An IGERB protein variant of the invention comprises an amino acid sequence identical to SEQ ID NO: 3 except for having one or more variant amino acids selected from the group consisting of methionine at a position corresponding to amino acid position 143, cysteine at a position corresponding to amino acid position 158, and alanine at a position corresponding to amino acid position 178, and may also comprise an additional variant amino acid of glycine at a position corresponding to amino acid position 237. The invention specifically excludes amino acid sequences identical to those previously identified for IGERB, including SEQ ID NO: 3, and previously described fragments thereof. IGERB protein variants included within the invention comprise all amino acid sequences based on SEQ ID NO: 3 and having the combination of amino acid variations described in Table 2 below. In preferred embodiments, an IGERB protein variant of the invention is encoded by an isogene defined by one of the observed haplotypes shown in Table 5.

Table2. Novel Polymorphic Variant of IGERB

Polymorphic Variant Number	Amino Acid Position and Identities			
	143	158	178	237
1	T	Y	A	E
2	T	Y	A	G
3	T	C	S	E
4	T	C	S	G
5	T	C	A	E
6	T	C	A	G
7	M	Y	S	E
8	M	Y	S	G
9	M	Y	A	E
10	M	Y	A	G
11	M	C	S	E
12	M	C	S	G
13	M	C	A	E
14	M	C	A	G

The invention also includes IGERB peptide variants, which are any fragments of an IGERB protein variant that contains one or more of the amino acid variations shown in Table 2. An IGERB peptide variant is at least 6 amino acids in length and is preferably any number between 6 and 30 amino acids long, more preferably between 10 and 25, and most preferably between 15 and 20 amino acids long. Such IGERB peptide variants may be useful as antigens to generate antibodies specific for one of the above IGERB isoforms. In addition, the IGERB peptide variants may be useful in drug screening assays.

An IGERB variant protein or peptide of the invention may be prepared by chemical synthesis or by expressing one of the variant IGERB genomic and cDNA sequences as described above. Alternatively, the IGERB protein variant may be isolated from a biological sample of an individual having an IGERB isogene which encodes the variant protein. Where the sample contains two different IGERB isoforms (i.e., the individual has different IGERB isogenes), a particular IGERB isoform of the invention can be isolated by immunoaffinity chromatography using an antibody which specifically binds to that particular IGERB isoform but does not bind to the other IGERB isoform.

The expressed or isolated IGERB protein may be detected by methods known in the art, including Coomassie blue staining, silver staining, and Western blot analysis using antibodies specific for the isoform of the IGERB protein as discussed further below. IGERB variant proteins can be purified by standard protein purification procedures known in the art, including differential precipitation, molecular sieve chromatography, ion-exchange chromatography, isoelectric focusing, gel electrophoresis, affinity and immunoaffinity chromatography and the like. (Ausubel et. al., 1987, In Current Protocols in Molecular Biology John Wiley and Sons, New York, New York). In the case of immunoaffinity chromatography, antibodies specific for a particular polymorphic variant may be used.

A polymorphic variant IGERB gene of the invention may also be fused in frame with a heterologous sequence to encode a chimeric IGERB protein. The non-IGERB portion of the chimeric

protein may be recognized by a commercially available antibody. In addition, the chimeric protein may also be engineered to contain a cleavage site located between the IGERB and non-IGERB portions so that the IGERB protein may be cleaved and purified away from the non-IGERB portion.

An additional embodiment of the invention relates to using a novel IGERB protein isoform in any of a variety of drug screening assays. Such screening assays may be performed to identify agents that bind specifically to all known IGERB protein isoforms or to only a subset of one or more of these isoforms. The agents may be from chemical compound libraries, peptide libraries and the like. The IGERB protein or peptide variant may be free in solution or affixed to a solid support. In one embodiment, high throughput screening of compounds for binding to an IGERB variant may be accomplished using the method described in PCT application WO84/03565, in which large numbers of test compounds are synthesized on a solid substrate, such as plastic pins or some other surface, contacted with the IGERB protein(s) of interest and then washed. Bound IGERB protein(s) are then detected using methods well-known in the art.

In another embodiment, a novel IGERB protein isoform may be used in assays to measure the binding affinities of one or more candidate drugs targeting the IGERB protein.

In another embodiment, the invention provides antibodies specific for and immunoreactive with one or more of the novel IGERB variant proteins described herein. The antibodies may be either monoclonal or polyclonal in origin. The IGERB protein or peptide variant used to generate the antibodies may be from natural or recombinant sources or produced by chemical synthesis using synthesis techniques known in the art. If the IGERB protein variant is of insufficient size to be antigenic, it may be conjugated, complexed, or otherwise covalently linked to a carrier molecule to enhance the antigenicity of the peptide. Examples of carrier molecules, include, but are not limited to, albumins (e.g., human, bovine, fish, ovine), and keyhole limpet hemocyanin (Basic and Clinical Immunology, 1991, Eds. D.P. Stites, and A.I. Terr, Appleton and Lange, Norwalk Connecticut, San Mateo, California).

In one embodiment, an antibody specifically immunoreactive with one of the novel IGERB protein isoforms described herein is administered to an individual to neutralize activity of the IGERB isoform expressed by that individual. The antibody may be formulated as a pharmaceutical composition which includes a pharmaceutically acceptable carrier.

Antibodies specific for and immunoreactive with one of the novel IGERB protein isoforms described herein may be used to immunoprecipitate the IGERB protein variant from solution as well as react with IGERB protein isoforms on Western or immunoblots of polyacrylamide gels on membrane supports or substrates. In another preferred embodiment, the antibodies will detect IGERB protein isoforms in paraffin or frozen tissue sections, or in cells which have been fixed or unfixed and prepared on slides, coverslips, or the like, for use in immunocytochemical, immunohistochemical, and immunofluorescence techniques.

In another embodiment, an antibody specifically immunoreactive with one of the novel IGERB

protein variants described herein is used in immunoassays to detect this variant in biological samples. In this method, an antibody of the present invention is contacted with a biological sample and the formation of a complex between the IGERB protein variant and the antibody is detected. As described, suitable immunoassays include radioimmunoassay, Western blot assay, immunofluorescent assay, enzyme linked immunoassay (ELISA), chemiluminescent assay, immunohistochemical assay, immunocytochemical assay, and the like (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Press, New York, New York; Current Protocols in Molecular Biology, 1987, Eds. Ausubel et al., John Wiley and Sons, New York, New York). Standard techniques known in the art for ELISA are described in Methods in Immunodiagnosis, 2nd Ed., Eds. Rose and Bigazzi, John Wiley and Sons, New York 1980; and Campbell et al., 1984, Methods in Immunology, W.A. Benjamin, Inc.). Such assays may be direct, indirect, competitive, or noncompetitive as described in the art (see, e.g., Principles and Practice of Immunoassay, 1991, Eds. Christopher P. Price and David J. Neoman, Stockton Pres, NY, NY; and Oellirich, M., 1984, J. Clin. Chem. Clin. Biochem., 22:895-904). Proteins may be isolated from test specimens and biological samples by conventional methods, as described in Current Protocols in Molecular Biology, supra.

Exemplary antibody molecules for use in the detection and therapy methods of the present invention are intact immunoglobulin molecules, substantially intact immunoglobulin molecules, or those portions of immunoglobulin molecules that contain the antigen binding site. Polyclonal or monoclonal antibodies may be produced by methods conventionally known in the art (e.g., Kohler and Milstein, 1975, Nature, 256:495-497; Campbell Monoclonal Antibody Technology, the Production and Characterization of Rodent and Human Hybridomas, 1985, In: Laboratory Techniques in Biochemistry and Molecular Biology, Eds. Burdon et al., Volume 13, Elsevier Science Publishers, Amsterdam). The antibodies or antigen binding fragments thereof may also be produced by genetic engineering. The technology for expression of both heavy and light chain genes in E. coli is the subject of PCT patent applications, publication number WO 901443, WO 901443 and WO 9014424 and in Huse et al., 1989, Science, 246:1275-1281. The antibodies may also be humanized (e.g., Queen, C. et al. 1989 Proc. Natl. Acad. Sci. 86;10029).

Effect(s) of the polymorphisms identified herein on expression of IGERB may be investigated by preparing recombinant cells and/or organisms, preferably recombinant animals, containing a polymorphic variant of the IGERB gene. As used herein, "expression" includes but is not limited to one or more of the following: transcription of the gene into precursor mRNA; splicing and other processing of the precursor mRNA to produce mature mRNA; mRNA stability; translation of the mature mRNA into IGERB protein (including codon usage and tRNA availability); and glycosylation and/or other modifications of the translation product, if required for proper expression and function.

To prepare a recombinant cell of the invention, the desired IGERB isogene may be introduced into the cell in a vector such that the isogene remains extrachromosomal. In such a situation, the gene will be expressed by the cell from the extrachromosomal location. In a preferred embodiment, the

IGERB isogene is introduced into a cell in such a way that it recombines with the endogenous IGERB gene present in the cell. Such recombination requires the occurrence of a double recombination event, thereby resulting in the desired IGERB gene polymorphism. Vectors for the introduction of genes both for recombination and for extrachromosomal maintenance are known in the art, and any suitable vector or vector construct may be used in the invention. Methods such as electroporation, particle bombardment, calcium phosphate co-precipitation and viral transduction for introducing DNA into cells are known in the art; therefore, the choice of method may lie with the competence and preference of the skilled practitioner. Examples of cells into which the IGERB isogene may be introduced include, but are not limited to, continuous culture cells, such as COS, NIH/3T3, and primary or culture cells of the relevant tissue type, i.e., they express the IGERB isogene. Such recombinant cells can be used to compare the biological activities of the different protein variants.

Recombinant organisms, i.e., transgenic animals, expressing a variant IGERB gene are prepared using standard procedures known in the art. Preferably, a construct comprising the variant gene is introduced into a nonhuman animal or an ancestor of the animal at an embryonic stage, i.e., the one-cell stage, or generally not later than about the eight-cell stage. Transgenic animals carrying the constructs of the invention can be made by several methods known to those having skill in the art. One method involves transfecting into the embryo a retrovirus constructed to contain one or more insulator elements, a gene or genes of interest, and other components known to those skilled in the art to provide a complete shuttle vector harboring the insulated gene(s) as a transgene, see e.g., U.S. Patent No. 5,610,053. Another method involves directly injecting a transgene into the embryo. A third method involves the use of embryonic stem cells. Examples of animals into which the IGERB isogenes may be introduced include, but are not limited to, mice, rats, other rodents, and nonhuman primates (see "The Introduction of Foreign Genes into Mice" and the cited references therein, In: Recombinant DNA, Eds. J.D. Watson, M. Gilman, J. Witkowski, and M. Zoller; W.H. Freeman and Company, New York, pages 254-272). Transgenic animals stably expressing a human IGERB isogene and producing human IGERB protein can be used as biological models for studying diseases related to abnormal IGERB expression and/or activity, and for screening and assaying various candidate drugs, compounds, and treatment regimens to reduce the symptoms or effects of these diseases.

An additional embodiment of the invention relates to pharmaceutical compositions for treating disorders affected by expression or function of a novel IGERB isogene described herein. The pharmaceutical composition may comprise any of the following active ingredients: a polynucleotide comprising one of these novel IGERB isogenes; an antisense oligonucleotide directed against one of the novel IGERB isogenes, a polynucleotide encoding such an antisense oligonucleotide, or another compound which inhibits expression of a novel IGERB isogene described herein. Preferably, the composition contains the active ingredient in a therapeutically effective amount. By therapeutically effective amount is meant that one or more of the symptoms relating to disorders affected by expression or function of a novel IGERB isogene is reduced and/or eliminated. The composition also comprises a

pharmaceutically acceptable carrier, examples of which include, but are not limited to, saline, buffered saline, dextrose, and water. Those skilled in the art may employ a formulation most suitable for the active ingredient, whether it is a polynucleotide, oligonucleotide, protein, peptide or small molecule antagonist. The pharmaceutical composition may be administered alone or in combination with at least one other agent, such as a stabilizing compound. Administration of the pharmaceutical composition may be by any number of routes including, but not limited to oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, intradermal, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal. Further details on techniques for formulation and administration may be found in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing Co., Easton, PA).

For any composition, determination of the therapeutically effective dose of active ingredient and/or the appropriate route of administration is well within the capability of those skilled in the art. For example, the dose can be estimated initially either in cell culture assays or in animal models. The animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans. The exact dosage will be determined by the practitioner, in light of factors relating to the patient requiring treatment, including but not limited to severity of the disease state, general health, age, weight and gender of the patient, diet, time and frequency of administration, other drugs being taken by the patient, and tolerance/response to the treatment.

Information on the identity of genotypes and haplotypes for the IGERB gene of any particular individual as well as the frequency of such genotypes and haplotypes in any particular population of individuals is expected to be useful for a variety of basic research and clinical applications. Thus, the invention also provides compositions and methods for detecting the novel IGERB polymorphisms identified herein.

The compositions comprise at least one IGERB genotyping oligonucleotide. In one embodiment, an IGERB genotyping oligonucleotide is a probe or primer capable of hybridizing to a target region that is located close to, or that contains, one of the novel polymorphic sites described herein. As used herein, the term "oligonucleotide" refers to a polynucleotide molecule having less than about 100 nucleotides. A preferred oligonucleotide of the invention is 10 to 35 nucleotides long. More preferably, the oligonucleotide is between 15 and 30, and most preferably, between 20 and 25 nucleotides in length. The oligonucleotide may be comprised of any phosphorylation state of ribonucleotides, deoxyribonucleotides, and acyclic nucleotide derivatives, and other functionally equivalent derivatives. Alternatively, oligonucleotides may have a phosphate-free backbone, which may be comprised of linkages such as carboxymethyl, acetamidate, carbamate, polyamide (peptide nucleic acid (PNA)) and the like (Varma, R. in Molecular Biology and Biotechnology, A Comprehensive Desk Reference, Ed. R. Meyers, VCH Publishers, Inc. (1995), pages 617-620). Oligonucleotides of the invention may be prepared by chemical synthesis using any suitable

methodology known in the art, or may be derived from a biological sample, for example, by restriction digestion. The oligonucleotides may be labeled, according to any technique known in the art, including use of radiolabels, fluorescent labels, enzymatic labels, proteins, haptens, antibodies, sequence tags and the like.

Genotyping oligonucleotides of the invention must be capable of specifically hybridizing to a target region of an IGERB polynucleotide, i.e., an IGERB isogene. As used herein, specific hybridization means the oligonucleotide forms an anti-parallel double-stranded structure with the target region under certain hybridizing conditions, while failing to form such a structure when incubated with a non-target region or a non-IGERB polynucleotide under the same hybridizing conditions. Preferably, the oligonucleotide specifically hybridizes to the target region under conventional high stringency conditions. The skilled artisan can readily design and test oligonucleotide probes and primers suitable for detecting polymorphisms in the IGERB gene using the polymorphism information provided herein in conjunction with the known sequence information for the IGERB gene and routine techniques.

A nucleic acid molecule such as an oligonucleotide or polynucleotide is said to be a "perfect" or "complete" complement of another nucleic acid molecule if every nucleotide of one of the molecules is complementary to the nucleotide at the corresponding position of the other molecule. A nucleic acid molecule is "substantially complementary" to another molecule if it hybridizes to that molecule with sufficient stability to remain in a duplex form under conventional low-stringency conditions. Conventional hybridization conditions are described, for example, by Sambrook J. et al., in Molecular Cloning, A Laboratory Manual, 2nd Edition, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989) and by Haymes, B.D. et al. in Nucleic Acid Hybridization, A Practical Approach, IRL Press, Washington, D.C. (1985). While perfectly complementary oligonucleotides are preferred for detecting polymorphisms, departures from complete complementarity are contemplated where such departures do not prevent the molecule from specifically hybridizing to the target region. For example, an oligonucleotide primer may have a non-complementary fragment at its 5′ end, with the remainder of the primer being complementary to the target region. Alternatively, non-complementary nucleotides may be interspersed into the oligonucleotide probe or primer as long as the resulting probe or primer is still capable of specifically hybridizing to the target region.

Preferred genotyping oligonucleotides of the invention are allele-specific oligonucleotides. As used herein, the term allele-specific oligonucleotide (ASO) means an oligonucleotide that is able, under sufficiently stringent conditions, to hybridize specifically to one allele of a gene, or other locus, at a target region containing a polymorphic site while not hybridizing to the corresponding region in another allele(s). As understood by the skilled artisan, allele-specificity will depend upon a variety of readily optimized stringency conditions, including salt and formamide concentrations, as well as temperatures for both the hybridization and washing steps. Examples of hybridization and washing conditions typically used for ASO probes are found in Kogan et al., "Genetic Prediction of Hemophilia A" in PCR Protocols, A Guide to Methods and Applications, Academic Press, 1990 and Ruano et al., 87 Proc. Natl.

Acad. Sci. USA 6296-6300, 1990. Typically, an allele-specific oligonucleotide will be perfectly complementary to one allele while containing a single mismatch for another allele.

Allele-specific oligonucleotide probes which usually provide good discrimination between different alleles are those in which a central position of the oligonucleotide probe aligns with the polymorphic site in the target region (e.g., approximately the 7th or 8th position in a 15 mer, the 8th or 9th position in a 16mer, the 10th or 11th position in a 20 mer). A preferred ASO probe for detecting IGERB gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

Accession No.: M89796.1

ATCACAAGTAAAAGC (SEQ ID NO:4) and its complement,
ATCACAAATAAAAGC (SEQ ID NO:5) and its complement,
CCTCATCCCCACCAC (SEQ ID NO:6) and its complement,
CCTCATCTCCACCAC (SEQ ID NO:7) and its complement,
GTATTAAGATGATAT (SEQ ID NO:8) and its complement,
GTATTAACATGATAT (SEQ ID NO:9) and its complement,
TATAACATAGATATG (SEQ ID NO:10) and its complement,
TATAACACAGATATG (SEQ ID NO:11) and its complement,
GGGGGAACGGGAATT (SEQ ID NO:12) and its complement,
GGGGGAATGGGAATT (SEQ ID NO:13) and its complement,
TTGGCCTATATCCAC (SEQ ID NO:14) and its complement,
TTGGCCTGTATCCAC (SEQ ID NO:15) and its complement,
TTCCTTTTCCACTGT (SEQ ID NO:16) and its complement,
TTCCTTTGCCACTGT (SEQ ID NO:17) and its complement,
TTTTTTTTTGTGTGG (SEQ ID NO:18) and its complement,
TTTTTTTGTGTGTGG (SEQ ID NO:19) and its complement,
TTTTTGTGTGGGAAG (SEQ ID NO:20) and its complement,
TTTTTGTATGGGAAG (SEQ ID NO:21) and its complement,
GGAAAAATTTCTATT (SEQ ID NO:22) and its complement,
GGAAAAACTTCTATT (SEQ ID NO:23) and its complement,
ATTTATTCGCCTGAT (SEQ ID NO:24) and its complement, and
ATTTATTTGCCTGAT (SEQ ID NO:25) and its complement.

An allele-specific oligonucleotide primer of the invention has a 3′ terminal nucleotide, or preferably a 3′ penultimate nucleotide, that is complementary to only one nucleotide of a particular SNP, thereby acting as a primer for polymerase-mediated extension only if the allele containing that nucleotide is present. Allele-specific oligonucleotide primers hybridizing to either the coding or noncoding strand are contemplated by the invention. A preferred ASO primer for detecting IGERB gene polymorphisms comprises a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

Accession No.:M89796.1

CTATTCATCACAAGT (SEQ ID NO:26); CAACAGGCTTTTACT (SEQ ID NO:27);
CTATTCATCACAAAT (SEQ ID NO:28); CAACAGGCTTTTATT (SEQ ID NO:29);
AGTCGGCCTCATCCC (SEQ ID NO:30); TATGCAGTGGTGGGG (SEQ ID NO:31);
AGTCGGCCTCATCTC (SEQ ID NO:32); TATGCAGTGGTGGAG (SEQ ID NO:33);

TTCCCAGTATTAAGA (SEQ ID NO:34); TTATAAATATCATCT (SEQ ID NO:35);
TTCCCAGTATTAACA (SEQ ID NO:36); TTATAAATATCATGT (SEQ ID NO:37);
AGCATATATAACATA (SEQ ID NO:38); AATGAGCATATCTAT (SEQ ID NO:39);
AGCATATATAACACA (SEQ ID NO:40); AATGAGCATATCTGT (SEQ ID NO:41);
ATAGCTGGGGGAACG (SEQ ID NO:42); GATGGTAATTCCCGT (SEQ ID NO:43);
ATAGCTGGGGGAATG (SEQ ID NO:44); GATGGTAATTCCCAT (SEQ ID NO:45);
AAGAGCTTGGCCTAT (SEQ ID NO:46); GTGGATGTGGATATA (SEQ ID NO:47);
AAGAGCTTGGCCTGT (SEQ ID NO:48); GTGGATGTGGATACA (SEQ ID NO:49);
TATGGCTTCCTTTTC (SEQ ID NO:50); ATACATACAGTGGAA (SEQ ID NO:51);
TATGGCTTCCTTTGC (SEQ ID NO:52); ATACATACAGTGGCA (SEQ ID NO:53);
TATGTATTTTTTTTT (SEQ ID NO:54); GTCTTCCCACACAAA (SEQ ID NO:55);
TATGTATTTTTTTGT (SEQ ID NO:56); GTCTTCCCACACACA (SEQ ID NO:57);
TATTTTTTTTTGTGT (SEQ ID NO:58); CTTAGTCTTCCCACA (SEQ ID NO:59);
TATTTTTTTTTGTAT (SEQ ID NO:60); CTTAGTCTTCCCATA (SEQ ID NO:61);
GCTACTGGAAAAATT (SEQ ID NO:62); GGAGAGAATAGAAAT (SEQ ID NO:63);
GCTACTGGAAAAACT (SEQ ID NO:64); GGAGAGAATAGAAGT (SEQ ID NO:65);
CATTAGATTTATTCG (SEQ ID NO:66); ATTCTTATCAGGCGA (SEQ ID NO:67);
CATTAGATTTATTTG (SEQ ID NO:68); and
ATTCTTATCAGGCAA (SEQ ID NO:69).

Other genotyping oligonucleotides of the invention hybridize to a target region located one to several nucleotides downstream of one of the novel polymorphic sites identified herein. Such oligonucleotides are useful in polymerase-mediated primer extension methods for detecting one of the novel polymorphisms described herein and therefore such genotyping oligonucleotides are referred to herein as "primer-extension oligonucleotides". In a preferred embodiment, the 3′-terminus of a primer-extension oligonucleotide is a deoxynucleotide complementary to the nucleotide located immediately adjacent to the polymorphic site. A particularly preferred oligonucleotide primer for detecting IGERB gene polymorphisms by primer extension terminates in a nucleotide sequence, listed 5′ to 3′, selected from the group consisting of:

Accession No.:M89796.1

TTCATCACAA (SEQ ID NO:70); CAGGCTTTTA (SEQ ID NO:71);
CGGCCTCATC (SEQ ID NO:72); GCAGTGGTGG (SEQ ID NO:73);
CCAGTATTAA (SEQ ID NO:74); TAAATATCAT (SEQ ID NO:75);
ATATATAACA (SEQ ID NO:76); GAGCATATCT (SEQ ID NO:77);
GCTGGGGGAA (SEQ ID NO:78); GGTAATTCCC (SEQ ID NO:79);
AGCTTGGCCT (SEQ ID NO:80); GATGTGGATA (SEQ ID NO:81);
GGCTTCCTTT (SEQ ID NO:82); CATACAGTGG (SEQ ID NO:83);
GTATTTTTTT (SEQ ID NO:84); TTCCCACACA (SEQ ID NO:85);
TTTTTTTTGT (SEQ ID NO:86); AGTCTTCCCA (SEQ ID NO:87);
ACTGGAAAAA (SEQ ID NO:88); GAGAATAGAA (SEQ ID NO:89);
TAGATTTATT (SEQ ID NO:90); and CTTATCAGGC (SEQ ID NO:91).

In some embodiments, a composition contains two or more differently labeled genotyping oligonucleotides for simultaneously probing the identity of nucleotides at two or more polymorphic sites. It is also contemplated that primer compositions may contain two or more sets of allele-specific primer pairs to allow simultaneous targeting and amplification of two or more regions containing a polymorphic site.

IGERB genotyping oligonucleotides of the invention may also be immobilized on or synthesized on a solid surface such as a microchip, bead, or glass slide (see, e.g., WO 98/20020 and WO 98/20019). Such immobilized genotyping oligonucleotides may be used in a variety of polymorphism detection assays, including but not limited to probe hybridization and polymerase extension assays. Immobilized IGERB genotyping oligonucleotides of the invention may comprise an ordered array of oligonucleotides designed to rapidly screen a DNA sample for polymorphisms in multiple genes at the same time.

In another embodiment, the invention provides a kit comprising at least two genotyping oligonucleotides packaged in separate containers. The kit may also contain other components such as hybridization buffer (where the oligonucleotides are to be used as a probe) packaged in a separate container. Alternatively, where the oligonucleotides are to be used to amplify a target region, the kit may contain, packaged in separate containers, a polymerase and a reaction buffer optimized for primer extension mediated by the polymerase, such as PCR.

The above described oligonucleotide compositions and kits are useful in methods for genotyping and/or haplotyping the IGERB gene in an individual. As used herein, the terms "IGERB genotype" and "IGERB haplotype" mean the genotype or haplotype contains the nucleotide pair or nucleotide, respectively, that is present at one or more of the novel polymorphic sites described herein and may optionally also include the nucleotide pair or nucleotide present at one or more additional polymorphic sites in the IGERB gene. The additional polymorphic sites may be currently known polymorphic sites or sites that are subsequently discovered.

One embodiment of the genotyping method involves isolating from the individual a nucleic acid mixture comprising the two copies of the IGERB gene, or a fragment thereof, that are present in the individual, and determining the identity of the nucleotide pair at one or more of the polymorphic sites selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13 in the two copies to assign an IGERB genotype to the individual. As will be readily understood by the skilled artisan, the two "copies" of a gene in an individual may be the same allele or may be different alleles. In a preferred embodiment of the genotyping method, the identity of the nucleotide pair atone or more of the polymorphic sites selected from the group consisting of PS1 and PS11 is also determined. In a particularly preferred embodiment, the genotyping method comprises determining the identity of the nucleotide pair at each of PS1-13.

Typically, the nucleic acid mixture is isolated from a biological sample taken from the individual, such as a blood sample or tissue sample. Suitable tissue samples include whole blood, semen saliva, tears, urine, fecal material, sweat, buccal, skin and hair. The nucleic acid mixture may be comprised of genomic DNA, mRNA, or cDNA and, in the latter two cases, the biological sample must be obtained from an organ in which the IGERB gene is expressed. Furthermore it will be understood by the skilled artisan that mRNA or cDNA preparations would not be used to detect polymorphisms located in introns or in 5′ and 3′ nontranscribed regions. If an IGERB gene fragment is isolated, it must

contain the polymorphic site(s) to be genotyped.

One embodiment of the haplotyping method comprises isolating from the individual a nucleic acid molecule containing only one of the two copies of the IGERB gene, or a fragment thereof, that is present in the individual and determining in that copy the identity of the nucleotide at one or more of the polymorphic sites PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13 in that copy to assign an IGERB haplotype to the individual. The nucleic acid may be isolated using any method capable of separating the two copies of the IGERB gene or fragment such as one of the methods described above for preparing IGERB isogenes, with targeted *in vivo* cloning being the preferred approach. As will be readily appreciated by those skilled in the art, any individual clone will only provide haplotype information on one of the two IGERB gene copies present in an individual. If haplotype information is desired for the individual's other copy, additional IGERB clones will need to be examined. Typically, at least five clones should be examined to have more than a 90% probability of haplotyping both copies of the IGERB gene in an individual. In some embodiments, the haplotyping method also comprises identifying the nucleotide at one or more of the polymorphic sites PS1 and PS11. In a particularly preferred embodiment, the nucleotide at each of PS1-13 is identified.

In a preferred embodiment, an IGERB haplotype pair is determined for an individual by identifying the phased sequence of nucleotides at one or more of the polymorphic sites selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13 in each copy of the IGERB gene that is present in the individual. In a particularly preferred embodiment, the haplotyping method comprises identifying the phased sequence of nucleotides at each of PS1-13 in each copy of the IGERB gene. When haplotyping both copies of the gene, the identifying step is preferably performed with each copy of the gene being placed in separate containers. However, it is also envisioned that if the two copies are labeled with different tags, or are otherwise separately distinguishable or identifiable, it could be possible in some cases to perform the method in the same container. For example, if first and second copies of the gene are labeled with different first and second fluorescent dyes, respectively, and an allele-specific oligonucleotide labeled with yet a third different fluorescent dye is used to assay the polymorphic site(s), then detecting a combination of the first and third dyes would identify the polymorphism in the first gene copy while detecting a combination of the second and third dyes would identify the polymorphism in the second gene copy.

In both the genotyping and haplotyping methods, the identity of a nucleotide (or nucleotide pair) at a polymorphic site(s) may be determined by amplifying a target region(s) containing the polymorphic site(s) directly from one or both copies of the IGERB gene, or fragment thereof, and the sequence of the amplified region(s) determined by conventional methods. It will be readily appreciated by the skilled artisan that only one nucleotide will be detected at a polymorphic site in individuals who are homozygous at that site, while two different nucleotides will be detected if the individual is heterozygous for that site. The polymorphism may be identified directly, known as positive-type identification, or by inference, referred to as negative-type identification. For example, where a SNP is

known to be guanine and cytosine in a reference population, a site may be positively determined to be either guanine or cytosine for an individual homozygous at that site, or both guanine and cytosine, if the individual is heterozygous at that site. Alternatively, the site may be negatively determined to be not guanine (and thus cytosine/cytosine) or not cytosine (and thus guanine/guanine).

In addition, the identity of the allele(s) present at any of the novel polymorphic sites described herein may be indirectly determined by genotyping a polymorphic site not disclosed herein that is in linkage disequilibrium with the polymorphic site that is of interest. Two sites are said to be in linkage disequilibrium if the presence of a particular variant at one site enhances the predictability of another variant at the second site (Stevens, JC 1999, *Mol. Diag.* **4**: 309-17). Polymorphic sites in linkage disequilibrium with the presently disclosed polymorphic sites may be located in regions of the gene or in other genomic regions not examined herein. Genotyping of a polymorphic site in linkage disequilibrium with the novel polymorphic sites described herein may be performed by, but is not limited to, any of the above-mentioned methods for detecting the identity of the allele at a polymorphic site.

The target region(s) may be amplified using any oligonucleotide-directed amplification method, including but not limited to polymerase chain reaction (PCR) (U.S. Patent No. 4,965,188), ligase chain reaction (LCR) (Barany et al., Proc. Natl. Acad. Sci. USA 88:189-193, 1991; WO90/01069), and oligonucleotide ligation assay (OLA) (Landegren et al., Science 241:1077-1080, 1988). Oligonucleotides useful as primers or probes in such methods should specifically hybridize to a region of the nucleic acid that contains or is adjacent to the polymorphic site. Typically, the oligonucleotides are between 10 and 35 nucleotides in length and preferably, between 15 and 30 nucleotides in length. Most preferably, the oligonucleotides are 20 to 25 nucleotides long. The exact length of the oligonucleotide will depend on many factors that are routinely considered and practiced by the skilled artisan.

Other known nucleic acid amplification procedures may be used to amplify the target region including transcription-based amplification systems (U.S. Patent No. 5,130,238; EP 329,822; U.S. Patent No. 5,169,766, WO89/06700) and isothermal methods (Walker et al., *Proc. Natl. Acad. Sci. USA* **89**:392-396, 1992).

A polymorphism in the target region may also be assayed before or after amplification using one of several hybridization-based methods known in the art. Typically, allele-specific oligonucleotides are utilized in performing such methods. The allele-specific oligonucleotides may be used as differently labeled probe pairs, with one member of the pair showing a perfect match to one variant of a target sequence and the other member showing a perfect match to a different variant. In some embodiments, more than one polymorphic site may be detected at once using a set of allele-specific oligonucleotides or oligonucleotide pairs. Preferably, the members of the set have melting temperatures within 5°C, and more preferably within 2°C, of each other when hybridizing to each of the polymorphic sites being detected.

Hybridization of an allele-specific oligonucleotide to a target polynucleotide may be performed with both entities in solution, or such hybridization may be performed when either the oligonucleotide or the target polynucleotide is covalently or noncovalently affixed to a solid support. Attachment may be mediated, for example, by antibody-antigen interactions, poly-L-Lys, streptavidin or avidin-biotin, salt bridges, hydrophobic interactions, chemical linkages, UV cross-linking baking, etc. Allele-specific oligonucleotides may be synthesized directly on the solid support or attached to the solid support subsequent to synthesis. Solid-supports suitable for use in detection methods of the invention include substrates made of silicon, glass, plastic, paper and the like, which may be formed, for example, into wells (as in 96-well plates), slides, sheets, membranes, fibers, chips, dishes, and beads. The solid support may be treated, coated or derivatized to facilitate the immobilization of the allele-specific oligonucleotide or target nucleic acid.

The genotype or haplotype for the IGERB gene of an individual may also be determined by hybridization of a nucleic sample containing one or both copies of the gene to nucleic acid arrays and subarrays such as described in WO 95/11995. The arrays would contain a battery of allele-specific oligonucleotides representing each of the polymorphic sites to be included in the genotype or haplotype.

The identity of polymorphisms may also be determined using a mismatch detection technique, including but not limited to the RNase protection method using riboprobes (Winter et al., Proc. Natl. Acad. Sci. USA 82:7575, 1985; Meyers et al., Science 230:1242, 1985) and proteins which recognize nucleotide mismatches, such as the E. coli mutS protein (Modrich, P. Ann. Rev. Genet. 25:229-253, 1991). Alternatively, variant alleles can be identified by single strand conformation polymorphism (SSCP) analysis (Orita et al., Genomics 5:874-879, 1989; Humphries et al., in Molecular Diagnosis of Genetic Diseases, R. Elles, ed., pp. 321-340, 1996) or denaturing gradient gel electrophoresis (DGGE) (Wartell et al., Nucl. Acids Res. 18:2699-2706, 1990; Sheffield et al., Proc. Natl. Acad. Sci. USA 86:232-236, 1989).

A polymerase-mediated primer extension method may also be used to identify the polymorphism(s). Several such methods have been described in the patent and scientific literature and include the "Genetic Bit Analysis" method (WO92/15712) and the ligase/polymerase mediated genetic bit analysis (U.S. Patent 5,679,524. Related methods are disclosed in WO91/02087, WO90/09455, WO95/17676, U.S. Patent Nos. 5,302,509, and 5,945,283. Extended primers containing a polymorphism may be detected by mass spectrometry as described in U.S. Patent No. 5,605,798. Another primer extension method is allele-specific PCR (Ruaño et al., Nucl. Acids Res. 17:8392, 1989; Ruaño et al., Nucl. Acids Res. 19, 6877-6882, 1991; WO 93/22456; Turki et al., J. Clin. Invest. 95:1635-1641, 1995). In addition, multiple polymorphic sites may be investigated by simultaneously amplifying multiple regions of the nucleic acid using sets of allele-specific primers as described in Wallace et al. (WO89/10414).

In another aspect of the invention, an individual's IGERB haplotype pair is predicted from its IGERB genotype using information on haplotype pairs known to exist in a reference population. In its

broadest embodiment, the haplotyping prediction method comprises identifying an IGERB genotype for the individual at two or more polymorphic sites selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13, enumerating all possible haplotype pairs which are consistent with the genotype, accessing data containing IGERB haplotype pairs identified in a reference population, and assigning a haplotype pair to the individual that is consistent with the data. In one embodiment, the reference haplotype pairs include the IGERB haplotype pairs shown in Table 4.

Generally, the reference population should be composed of randomly-selected individuals representing the major ethnogeographic groups of the world. A preferred reference population for use in the methods of the present invention comprises an approximately equal number of individuals from Caucasian, African American, Asian and Hispanic-Latino population groups with the minimum number of each group being chosen based on how rare a haplotype one wants to be guaranteed to see. For example, if one wants to have a q% chance of not missing a haplotype that exists in the population at a p% frequency of occurring in the reference population, the number of individuals (n) who must be sampled is given by $2n=\log(1-q)/\log(1-p)$ where p and q are expressed as fractions. A preferred reference population allows the detection of any haplotype whose frequency is at least 10% with about 99% certainty and comprises about 20 unrelated individuals from each of the four population groups named above. A particularly preferred reference population includes a 3-generation family representing one or more of the four population groups to serve as controls for checking quality of haplotyping procedures.

In a preferred embodiment, the haplotype frequency data for each ethnogeographic group is examined to determine whether it is consistent with Hardy-Weinberg equilibrium. Hardy-Weinberg equilibrium (D.L. Hartl et al., Principles of Population Genomics, Sinauer Associates (Sunderland, MA), 3rd Ed., 1997) postulates that the frequency of finding the haplotype pair H_1 / H_2 is equal to $p_{H-W}(H_1 / H_2) = 2p(H_1)p(H_2)$ if $H_1 \neq H_2$ and $p_{H-W}(H_1 / H_2) = p(H_1)p(H_2)$ if $H_1 = H_2$. A statistically significant difference between the observed and expected haplotype frequencies could be due to one or more factors including significant inbreeding in the population group, strong selective pressure on the gene, sampling bias, and/or errors in the genotyping process. If large deviations from Hardy-Weinberg equilibrium are observed in an ethnogeographic group, the number of individuals in that group can be increased to see if the deviation is due to a sampling bias. If a larger sample size does not reduce the difference between observed and expected haplotype pair frequencies, then one may wish to consider haplotyping the individual using a direct haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843, 1996).

In one embodiment of this method for predicting an IGERB haplotype pair, the assigning step involves performing the following analysis. First, each of the possible haplotype pairs is compared to the haplotype pairs in the reference population. Generally, only one of the haplotype pairs in the reference population matches a possible haplotype pair and that pair is assigned to the individual.

Occasionally, only one haplotype represented in the reference haplotype pairs is consistent with a possible haplotype pair for an individual, and in such cases the individual is assigned a haplotype pair containing this known haplotype and a new haplotype derived by subtracting the known haplotype from the possible haplotype pair. In rare cases, either no haplotypes in the reference population are consistent with the possible haplotype pairs, or alternatively, multiple reference haplotype pairs are consistent with the possible haplotype pairs. In such cases, the individual is preferably haplotyped using a direct molecular haplotyping method such as, for example, CLASPER System™ technology (U.S. Patent No. 5,866,404), SMD, or allele-specific long-range PCR (Michalotos-Beloin et al., Nucleic Acids Res. 24:4841-4843, 1996). A preferred process for predicting IGERB haplotype pairs from IGERB genotypes is described in copending U.S. Provisional Application Serial No. 60/198,340.

The invention also provides a method for determining the frequency of an IGERB genotype or IGERB haplotype in a population. The method comprises determining the genotype or the haplotype pair for the IGERB gene that is present in each member of the population, wherein the genotype or haplotype comprises the nucleotide pair or nucleotide detected at one or more of the polymorphic sites PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13 in the IGERB gene; and calculating the frequency any particular genotype or haplotype is found in the population. The population may be a reference population, a family population, a same sex population, a population group, a trait population (e.g., a group of individuals exhibiting a trait of interest such as a medical condition or response to a therapeutic treatment).

In another aspect of the invention, frequency data for IGERB genotypes and/or haplotypes found in a reference population are used in a method for identifying an association between a trait and an IGERB genotype or an IGERB haplotype. The trait may be any detectable phenotype, including but not limited to susceptibility to a disease or response to a treatment. The method involves obtaining data on the frequency of the genotype(s) or haplotype(s) of interest in a reference population as well as in a population exhibiting the trait. Frequency data for one or both of the reference and trait populations may be obtained by genotyping or haplotyping each individual in the populations using one of the methods described above. The haplotypes for the trait population may be determined directly or, alternatively, by the predictive genotype to haplotype approach described above. In another embodiment, the frequency data for the reference and/or trait populations is obtained by accessing previously determined frequency data, which may be in written or electronic form. For example, the frequency data may be present in a database that is accessible by a computer. Once the frequency data is obtained, the frequencies of the genotype(s) or haplotype(s) of interest in the reference and trait populations are compared. In a preferred embodiment, the frequencies of all genotypes and/or haplotypes observed in the populations are compared. If a particular genotype or haplotype for the IGERB gene is more frequent in the trait population than in the reference population at a statistically significant amount, then the trait is predicted to be associated with that IGERB genotype or haplotype. Preferably, the IGERB genotype or haplotype being compared in the trait and reference populations is

selected from the full-genotypes and full-haplotypes shown in Tables 4 and 5, respectively, or from sub-genotypes and sub-haplotypes derived from these genotypes and haplotypes.

In a preferred embodiment of the method, the trait of interest is a clinical response exhibited by a patient to some therapeutic treatment, for example, response to a drug targeting IGERB or response to a therapeutic treatment for a medical condition. As used herein, "medical condition" includes but is not limited to any condition or disease manifested as one or more physical and/or psychological symptoms for which treatment is desirable, and includes previously and newly identified diseases and other disorders. As used herein the term "clinical response" means any or all of the following: a quantitative measure of the response, no response, and adverse response (i.e., side effects).

In order to deduce a correlation between clinical response to a treatment and an IGERB genotype or haplotype, it is necessary to obtain data on the clinical responses exhibited by a population of individuals who received the treatment, hereinafter the "clinical population". This clinical data may be obtained by analyzing the results of a clinical trial that has already been run and/or the clinical data may be obtained by designing and carrying out one or more new clinical trials. As used herein, the term "clinical trial" means any research study designed to collect clinical data on responses to a particular treatment, and includes but is not limited to phase I, phase II and phase III clinical trials. Standard methods are used to define the patient population and to enroll subjects.

It is preferred that the individuals included in the clinical population have been graded for the existence of the medical condition of interest. This is important in cases where the symptom(s) being presented by the patients can be caused by more than one underlying condition, and where treatment of the underlying conditions are not the same. An example of this would be where patients experience breathing difficulties that are due to either asthma or respiratory infections. If both sets were treated with an asthma medication, there would be a spurious group of apparent non-responders that did not actually have asthma. These people would affect the ability to detect any correlation between haplotype and treatment outcome. This grading of potential patients could employ a standard physical exam or one or more lab tests. Alternatively, grading of patients could use haplotyping for situations where there is a strong correlation between haplotype pair and disease susceptibility or severity.

The therapeutic treatment of interest is administered to each individual in the trial population and each individual's response to the treatment is measured using one or more predetermined criteria. It is contemplated that in many cases, the trial population will exhibit a range of responses and that the investigator will choose the number of responder groups (e.g., low, medium, high) made up by the various responses. In addition, the IGERB gene for each individual in the trial population is genotyped and/or haplotyped, which may be done before or after administering the treatment.

After both the clinical and polymorphism data have been obtained, correlations between individual response and IGERB genotype or haplotype content are created. Correlations may be produced in several ways. In one method, individuals are grouped by their IGERB genotype or haplotype (or haplotype pair) (also referred to as a polymorphism group), and then the averages and

standard deviations of clinical responses exhibited by the members of each polymorphism group are calculated.

These results are then analyzed to determine if any observed variation in clinical response between polymorphism groups is statistically significant. Statistical analysis methods which may be used are described in L.D. Fisher and G. vanBelle, "Biostatistics: A Methodology for the Health Sciences", Wiley-Interscience (New York) 1993. This analysis may also include a regression calculation of which polymorphic sites in the IGERB gene give the most significant contribution to the differences in phenotype. One regression model useful in the invention is described in PCT Application Serial No. PCT/US00/17540, entitled "Methods for Obtaining and Using Haplotype Data".

A second method for finding correlations between IGERB haplotype content and clinical responses uses predictive models based on error-minimizing optimization algorithms. One of many possible optimization algorithms is a genetic algorithm (R. Judson, "Genetic Algorithms and Their Uses in Chemistry" in Reviews in Computational Chemistry, Vol. 10, pp. 1-73, K. B. Lipkowitz and D. B. Boyd, eds. (VCH Publishers, New York, 1997). Simulated annealing (Press et al., "Numerical Recipes in C: The Art of Scientific Computing", Cambridge University Press (Cambridge) 1992, Ch. 10), neural networks (E. Rich and K. Knight, "Artificial Intelligence", 2nd Edition (McGraw-Hill, New York, 1991, Ch. 18), standard gradient descent methods (Press et al., supra Ch. 10), or other global or local optimization approaches (see discussion in Judson, supra) could also be used. Preferably, the correlation is found using a genetic algorithm approach as described in PCT Application Serial No. PCT/US00/17540.

Correlations may also be analyzed using analysis of variation (ANOVA) techniques to determine how much of the variation in the clinical data is explained by different subsets of the polymorphic sites in the IGERB gene. As described in PCT Application Serial No. PCT/US00/17540, ANOVA is used to test hypotheses about whether a response variable is caused by or correlated with one or more traits or variables that can be measured (Fisher and vanBelle, supra, Ch. 10).

From the analyses described above, a mathematical model may be readily constructed by the skilled artisan that predicts clinical response as a function of IGERB genotype or haplotype content. Preferably, the model is validated in one or more follow-up clinical trials designed to test the model.

The identification of an association between a clinical response and a genotype or haplotype (or haplotype pair) for the IGERB gene may be the basis for designing a diagnostic method to determine those individuals who will or will not respond to the treatment, or alternatively, will respond at a lower level and thus may require more treatment, i.e., a greater dose of a drug. The diagnostic method may take one of several forms: for example, a direct DNA test (i.e., genotyping or haplotyping one or more of the polymorphic sites in the IGERB gene), a serological test, or a physical exam measurement. The only requirement is that there be a good correlation between the diagnostic test results and the underlying IGERB genotype or haplotype that is in turn correlated with the clinical response. In a preferred embodiment, this diagnostic method uses the predictive haplotyping method described above.

Any or all analytical and mathematical operations involved in practicing the methods of the present invention may be implemented by a computer. In addition, the computer may execute a program that generates views (or screens) displayed on a display device and with which the user can interact to view and analyze large amounts of information relating to the IGERB gene and its genomic variation, including chromosome location, gene structure, and gene family, gene expression data, polymorphism data, genetic sequence data, and clinical data population data (e.g., data on ethnogeographic origin, clinical responses, genotypes, and haplotypes for one or more populations). The IGERB polymorphism data described herein may be stored as part of a relational database (e.g., an instance of an Oracle database or a set of ASCII flat files). These polymorphism data may be stored on the computer's hard drive or may, for example, be stored on a CD ROM or on one or more other storage devices accessible by the computer. For example, the data may be stored on one or more databases in communication with the computer via a network.

Preferred embodiments of the invention are described in the following examples. Other embodiments within the scope of the claims herein will be apparent to one skilled in the art from consideration of the specification or practice of the invention as disclosed herein. It is intended that the specification, together with the examples, be considered exemplary only, with the scope and spirit of the invention being indicated by the claims which follow the examples.

EXAMPLES

The Examples herein are meant to exemplify the various aspects of carrying out the invention and are not intended to limit the scope of the invention in any way. The Examples do not include detailed descriptions for conventional methods employed, such as in the performance of genomic DNA isolation, PCR and sequencing procedures. Such methods are well-known to those skilled in the art and are described in numerous publications, for example, Sambrook, Fritsch, and Maniatis, "Molecular Cloning: A Laboratory Manual", 2nd Edition, Cold Spring Harbor Laboratory Press, USA, (1989).

Example 1A

This example illustrates examination of various regions of the IGERB gene for polymorphic sites using DNA for Index Repository 1A.

Amplification of Target Regions

The following target regions of the IGERB gene were amplified using the PCR primer pairs listed below, with the sequences presented in the 5′ to 3′ direction and nucleotide positions shown for each region corresponding to the indicated GenBank Accession No.

Accession Number: M89796.1
Fragment 1
Forward Primer

13-34 GGTGCAATTGGATAACTTCTGC (SEQ ID NO:92)
Reverse Primer
Complement of 583-560 CTAAGCACCGTGACTATGACTTCC (SEQ ID NO:93)
PCR product 571 nt

Fragment 2
Forward Primer
63-85 GTGGGGACAATTCCAGAAGAAGG (SEQ ID NO:94)
Reverse Primer
Complement of 768-744 CCTGTCTCATACCCAGAAAATGAGC (SEQ ID NO:95)
PCR product 706 nt

Fragment 3
Forward Primer
943-968 CCCATCTAGCTATTCAAAGCCATTCC (SEQ ID NO:96)
Reverse Primer
Complement of 1727-1706 TGTTGGCCATTCTCGTCCTAGC (SEQ ID NO:97)
PCR product 785 nt

Fragment 4
Forward Primer
1821-1844 AGTATTGGCCTTCCATGCATTAGG (SEQ ID NO:98)
Reverse Primer
Complement of 2491-2466 TGCTAGAAGTATGTTCCTGGAGTTGG (SEQ ID NO:99)
PCR product 671 nt

Fragment 5
Forward Primer
4877-4898 ATGTGGTTCCTGAAGGCAGTCC (SEQ ID NO:100)
Reverse Primer
Complement of 5467-5445 AAAACCCAGAGATCGTCACTTGC (SEQ ID NO:101)
PCR product 591 nt

Fragment 6
Forward Primer
5422-5446 CCTCGGGGTTAAAGTTATCTACTGC (SEQ ID NO:102)
Reverse Primer
Complement of 6025-6002 TCCTCACAAGCCTTCTGTACATCC (SEQ ID NO:103)
PCR product 604 nt

Fragment 7
Forward Primer
7124-7149 ACCCAGAAAGACAAAAGTAGATGAGG (SEQ ID NO:104)
Reverse Primer
Complement of 7569-7547 TTCCAGCAGAGATGTGTGTGGTC (SEQ ID NO:105)
PCR product 446 nt

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of Index Repository 1A. The PCR reactions were carried out under the following conditions:

Reaction volume	= 20 µl
10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 2 µl
100 ng of human genomic DNA	= 1 µl
10 mM dNTP	= 0.4 µl

Advantage 2 Polymerase enzyme mix (Clontech)	= 0.2 µl
Forward Primer (10 µM)	= 0.4 µl
Reverse Primer (10 µM)	= 0.4 µl
Water	=15.6µl

Amplification profile:
94°C - 2 min. 1 cycle

94°C - 30 sec.
70°C - 45 sec. } 10 cycles
72°C - 1 min.

94°C - 30 sec.
64°C - 45 sec. } 35 cycles
72°C - 1 min.

Sequencing of PCR Products

The PCR products were purified by Solid Phase Reversible Immobilization using the protocol developed by the Whitehead Genome Center. A detailed protocol can be found at http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI_pcr.html.

Briefly, five µl of carboxyl coated magnetic beads (10 mg/ml) and 60 µl of HYB BUFFER (2.5M NaCl/20% PEG 8000) were added to each PCR reaction mixture (20 µl). The reaction mixture was mixed well and incubated at room temperature (RT) for 10 min. The microtitre plate was placed on a magnet for 2 min and the beads washed twice with 150 µl of 70% EtOH. The beads were air dried for 2 min and the DNA was eluted in 25 µl of distilled water and incubated at RT for 5 min. The beads were magnetically separated and the supernatant removed for testing and sequencing.

The purified PCR products were sequenced in both directions using the primer sets described previously or those listed, in the 5′ to 3′ direction, below.

Accession Number: M89796.1
Fragment 1
Forward Primer
52-71 TTGGGACACAAGTGGGGACA (SEQ ID NO:106)
Reverse Primer
Complement of 508-489 GAAGGCTCCTGTGGGAGAGC (SEQ ID NO:107)

Fragment 2
Forward Primer
164-184 TCCTGCTAGTCTCAGGCAAAA (SEQ ID NO:108)
Reverse Primer
Complement of 690-670 CTCCAGGGAGAGAGAACCACA (SEQ ID NO:109)

Fragment 3
Forward Primer
1149-1168 TTGGCATGGATTTTGCTCCA (SEQ ID NO:110)
Reverse Primer
Complement of 1682-1662 TGCTTCTGGCTCTTCCCAAAA (SEQ ID NO:111)

Fragment 4
Forward Primer
1895-1916 CGTGTGGATCATTTCTCAGGAC (SEQ ID NO:112)
Reverse Primer
Complement of 2434-2415 TCAAGTGGCCCTTAGGCAAG (SEQ ID NO:113)

Fragment 5
Forward Primer
4931-4951 TGCGGACATTTTCAGGGTTTC (SEQ ID NO:114)

Reverse Primer
Complement of 5386-5367 TCTCCCTGGGCTGTGTGAAC (SEQ ID NO:115)

Fragment 6
Forward Primer
5459-5479 CTGGGTTTTTCTGTGCCTGTG (SEQ ID NO:116)
Reverse Primer
Complement of 5949-5929 TCCCAACCCTCATTCAGAGGA (SEQ ID NO:117)

Fragment 7
Forward Primer
7151-7168 AAGTCTCTTGAGCGAGAC (SEQ ID NO:118)
Reverse Primer
Complement of 7488-7471 AAGGTGGACAGAAGCAGC (SEQ ID NO:119)

Analysis of Sequences for Polymorphic Sites

Sequences were analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., Nucleic Acids Res. 14:2745-2751, 1997). The presence of a polymorphism was confirmed on both strands. The polymorphisms and their locations in the IGERB gene are listed in Table 3 below.

Example 1B

This example illustrates examination of the IGERB gene for polymorphic sites in a region spanning about 260 base pairs upstream of the ATG start codon to about 260 base pairs downstream of the termination codon.

Amplification of Target Regions

PCR primer pairs, which were designed based on GenBank Accession No. M89796, are set forth below:

Fragment 1 (Promoter region/Exon 1)
Forward primer:
178-200 5′- GGCAAAATTATGCTCCAGGAGTC-3′ (SEQ ID NO:120)
Reverse primer:
581-559 5′- AAGCACCGTGACTATGACTTCCC-3′ (SEQ ID NO:121)
PCR product: 404

Fragment 2 (Exon 2)

Forward primer:		
1209-1232	5'- TTGGTCAGTTACTTGGATGCTCTG -3'	(SEQ ID NO:122)
Reverse primer:		
1676-1657	5'- TGGCTCTTCCCAAAAGGACC -3'	(SEQ ID NO:123)
PCR product: 468		
Fragment 3 (Exon 3)		
Forward primer:		
1895-1918	5'- CGTGTGGATCATTTCTCAGGACAG -3'	(SEQ ID NO:124)
Reverse primer:		
2387-2365	5'- TCTCATGGGATAAAACTGTGGGC -3'	(SEQ ID NO:125)
PCR product: 493		
Fragment 4 (Exon 4)		
Forward primer:		
4390-4413	5'- TTTACCTATGTTTGGAAGATGGGG -3'	(SEQ ID NO:126)
Reverse primer:		
4810-4787	5'- GAATATGCCTACAAGGACAATGCC -3'	(SEQ ID NO:127)
PCR product: 421		
Fragment 5 (Exon 5)		
Forward primer:		
4998-5020	5'- TCCAGCCCTGAAATGAAGATAGG -3'	(SEQ ID NO:128)
Reverse primer:		
5436-5414	5'- ACTTTAACCCCGAGGAATTTGC -3'	(SEQ ID NO:129)
PCR product: 439		
Fragment 6 (Exon 6)		
Forward primer:		
5526-5548	5'- AAAGGACTGGTCAGATGGTAGGG -3'	(SEQ ID NO:130)
Reverse primer:		
5998-5975	5'- TCTCATGCTCCACACACTTTAAGG -3'	(SEQ ID NO:131)
PCR product: 473		
Fragment 7 (Exon 7)		
Forward primer:		
7156-7179	5'- TCTTGAGCGAGACTTCTAGGGATG -3'	(SEQ ID NO:132)
Reverse primer:		
7579-7556	5'- CATGTTGACTTTCCAGCAGAGATG -3'	(SEQ ID NO:133)
PCR product: 424		

These primer pairs were used in PCR reactions containing genomic DNA isolated from immortalized cell lines for each member of a reference population of 70 human individuals. The PCR reactions were carried out under the following conditions:

Reaction volume	=50 µl
10 x Advantage 2 Polymerase reaction buffer (Clontech)	= 5 µl
100 ng of human genomic DNA	= 5 µl
10 mM dNTP	= 1 µl
Advantage 2 Polymerase enzyme mix (Clontech)	= 0.5 µl
Forward Primer (10 µM)	= 1 µl
Reverse Primer (10 µM)	= 1 µl

Water =36.5 µl

Amplification profile:
94°C - 2 min. 1 cycle

94°C - 30 sec.
70°C - 45 sec. } 10 cycles
72°C - 1 min.

94°C - 30 sec.
64°C - 45 sec. } 35 cycles
72°C - 1 min.

Sequencing of PCR Products

The PCR products were purified by Solid Phase Reversible Immobilization using the protocol developed by the Whitehead Genome Center. A detailed protocol can be found at http://www.genome.wi.mit.edu/sequencing/protocols/pure/SPRI_pcr.html.

Briefly, carboxyl coated magnetic beads (10 mg/ml) were washed three times with wash buffer (0.5 M EDTA, pH 8.0). Ten µl of washed beads and 50 µl of HYB BUFFER (2.5M NaCl/20% PEG 8000) were added to each PCR reaction mixture (50 µl). The reaction mixture was mixed well and incubated at RT for 10 min. The microtitre plate was placed on a magnet for 2 min and the beads washed twice with 150 µl of 70% EtOH. The beads were air dried for 2 min and resuspend in 20 µl of elution buffer (10 mM trisacetate, pH 7.8) and incubated at RT for 5 min. The beads were magnetically separated and the supernatant removed for testing and sequencing.

The purified PCR products were sequenced in both directions using the above PCR primer sets except where indicated below.

Fragment 1
Forward primer:
346-369 5′- CATGAGGTAACCCATTTCAACTGC-3′ (SEQ ID NO:134)
Reverse primer:
576-555 5′- CCGTGACTATGACTTCCCCTGC-3′ (SEQ ID NO:135)

Fragment 2
Forward primer:
1210-1234 5′- TGGTCAGTTACTTGGATGCTCTGAG-3′ (SEQ ID NO:136)
Reverse primer:
1670-1648 5′- TTCCCAAAAGGACCCAGTTAGTG-3′ (SEQ ID NO:137)

Fragment 4
Forward primer:
4401-4423 5′- TTGGAAGATGGGGTTAAAAGGAC -3′ (SEQ ID NO:138)
Reverse primer:
4805-4786 5′- TGCCTACAAGGACAATGCCG -3′ (SEQ ID NO:139)

Fragment 5

Forward primer:
5026-5046 5'- TGAATGTGCCAGCAAGTGCAG -3' (SEQ ID NO:140)
Reverse primer:
5422-5403 5'- GAATTTGCCTGGGTTGAGGG -3' (SEQ ID NO:141)

Fragment 6
Forward primer:
5534-5558 5'- GGTCAGATGGTAGGGAGATGAAAAC -3' (SEQ ID NO:142)
Reverse primer:
5996-5974 5'- TCATGCTCCACACACTTTAAGGC -3' (SEQ ID NO:143)

Fragment 7
Forward primer:
7159-7180 5'- TGAGCGAGACTTCTAGGGATGG -3' (SEQ ID NO:144)
Reverse primer:
7573-7550 5'- GACTTTCCAGCAGAGATGTGTGTG -3' (SEQ ID NO:145)

Analysis of Sequences for Polymorphic Sites

Sequences were analyzed for the presence of polymorphisms using the Polyphred program (Nickerson et al., *Nucleic Acids Res.* 14:2745-2751, 1997). The presence of a polymorphism was confirmed on both strands. The polymorphisms and their locations in the IgERB gene are listed in Table 3 below.

Table 3. Polymorphic Sites Identified in the IGERB Gene

Polymorphic Site Number	Nucleotide Position	Reference Allele	Variant Allele	Example
PS1	245(Acc#M89796.1)	C	T	1A, 1B
PS2	266(Acc#M89796.1)	G	A	1A
PS3	1456(Acc#M89796.1)	C	T	1A
PS4	2253(Acc#M89796.1)	G	C	1A, 1B
PS5	2302(Acc#M89796.1)	T	C	1A
PS6	5128(Acc#M89796.1)	C	T	1A
PS7	5173(Acc#M89796.1)	A	G	1A
PS8	5232(Acc#M89796.1)	T	G	1A
PS9	5252(Acc#M89796.1)	T	G	1B
PS10	5256(Acc#M89796.1)	G	A	1B
PS11	7297(Acc#M89796.1)	A	G	1A, 1B
PS12	7402(Acc#M89796.1)	T	C	1A
PS13	7446(Acc#M89796.1)	C	T	1A

Example 2

This example illustrates analysis of the IGERB polymorphisms identified in Index Repository 1A for human genotypes and haplotypes for all polymorphic sites except PS9 and PS10.

The different genotypes containing these polymorphisms that were observed in the reference population are shown in Table 4 below, with the haplotype pair indicating the combination of haplotypes determined for the individual using the haplotype derivation protocol described below. In Table 4, homozygous positions are indicated by one nucleotide and heterozygous positions are indicated by two nucleotides.

Table 4. Genotypes and Haplotype Pairs Observed for the IGERB Gene

Genotype Number	Polymorphic Sites											Hap Pair	
	PS1	PS2	PS3	PS4	PS5	PS6	PS7	PS8	PS11	PS12	PS13		
1	T	G	C	C	T	C	A	T	A	T	C	12	12
2	C	G	C	G	T	C	A	T	A	T	C	7	7
3	C	G	C	C	T	C	A	T	G	T	T	5	5
4	C/T	G	C	G/C	T	C	A	T	A	T/C	C	7	11
5	C/T	A/G	C	C	T/C	C/T	A	T	A	T	C	1	10
6	T	G	C	C	T/C	C/T	A	T	A	T	C	12	10
7	C	G	C	G	T	C	A/G	T	A	T	C	7	8
8	C	G	C/T	G	T	C	A	T	A	T	C	7	9
9	T/C	G	C	C	T	C	A	T	A/G	T	C/T	12	5
10	C	G	C	G/C	T	C	A	T/G	A/G	T	C/T	7	3
11	C	G	C	G/C	T	C	A	T	A/G	T	C/T	7	5
12	T/C	G	C	C	T	C	A	T	A/G	T	C	12	4
13	C	A/G	C	C	T	C	A	T	A/G	T	C	1	4
14	T/C	G/A	C	C	T	C	A	T	A	T	C	12	1
15	C	G/A	C	G/C	T	C	A	T	A	T	C	7	1
16	C	G	C	G/C	T	C	A	T	A/G	T	C	7	4
17	C	G	C	C	T/C	T	A	T	A	T	T	6	2
18	T/C	G	C	C/G	T	C	A	T	A	T	C	12	7

The haplotype pairs shown in Table 4 were estimated from the unphased genotypes using an extension of Clark's algorithm (Clark, A.G. (1990) *Mol Bio Evol* 7, 111-122), as described in U.S. Provisional Application Serial No. 60/198,340 entitled "A Method and System for Determining Haplotypes from a Collection of Polymorphisms". In this method, haplotypes are assigned directly from individuals who are homozygous at all sites or heterozygous at no more than one of the variable sites. This list of haplotypes is then used to deconvolute the unphased genotypes in the remaining (multiply heterozygous) individuals.

By following this protocol, it was determined that the Index Repository examined herein and, by extension, the general population contains the 12 human IGERB haplotypes shown in Table 5 below.

Table 5. Haplotypes Identified in the IGERB Gene

Haplotype Number	Polymorphic Sites										
	PS 1	PS 2	PS 3	PS 4	PS 5	PS 6	PS 7	PS 8	PS 11	PS 12	PS 13
1	C	A	C	C	T	C	A	T	A	T	C
2	C	G	C	C	C	T	A	T	A	T	T
3	C	G	C	C	T	C	A	G	G	T	T
4	C	G	C	C	T	C	A	T	G	T	C
5	C	G	C	C	T	C	A	T	G	T	T
6	C	G	C	C	T	T	A	T	A	T	T
7	C	G	C	G	T	C	A	T	A	T	C
8	C	G	C	G	T	C	G	T	A	T	C
9	C	G	T	G	T	C	A	T	A	T	C
10	T	G	C	C	C	T	A	T	A	T	C
11	T	G	C	C	T	C	A	T	A	C	C
12	T	G	C	C	T	C	A	T	A	T	C

In view of the above, it will be seen that the several advantages of the invention are achieved and other advantageous results attained.

As various changes could be made in the above methods and compositions without departing from the scope of the invention, it is intended that all matter contained in the above description and shown in the accompanying drawings shall be interpreted as illustrative and not in a limiting sense.

All references cited in this specification, including patents and patent applications, are hereby incorporated in their entirety by reference. The discussion of references herein is intended merely to summarize the assertions made by their authors and no admission is made that any reference constitutes prior art. Applicants reserve the right to challenge the accuracy and pertinency of the cited references.

What is Claimed is:

1. An isolated polynucleotide comprising a nucleotide sequence selected from the group consisting of:

(a) a first nucleotide sequence which is a polymorphic variant of a reference sequence for immunoglobulin E receptor beta chain (IGERB) gene or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:1, and the polymorphic variant comprises at least one polymorphism selected from the group consisting of adenine at PS2, thymine at PS3, cytosine at PS4, cytosine at PS5, thymine at PS6, guanine at PS7, guanine at PS8, guanine at PS9, adenine at PS10, cytosine at PS12 and thymine at PS13; and

(b) a second nucleotide sequence which is complementary to the first nucleotide sequence.

2. The isolated polynucleotide of claim 1 which comprises an IGERB isogene.

3. The isolated polynucleotide of claim 1 which is a DNA molecule and comprises both the first and second nucleotide sequences and further comprises expression regulatory elements operably linked to the first nucleotide sequence.

4. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 1, wherein the organism expresses an IGERB protein encoded by the first nucleotide sequence.

5. The recombinant organism of claim 4 which is a nonhuman transgenic animal.

6. The isolated polynucleotide of claim 1, wherein the first nucleotide sequence is a polymorphic variant of a fragment of the IGERB gene, the fragment comprising one or more polymorphisms selected from the group consisting of adenine at PS2, thymine at PS3, cytosine at PS4, cytosine at PS5, thymine at PS6, guanine at PS7, guanine at PS8, guanine at PS9, adenine at PS10, cytosine at PS12 and thymine at PS13.

7. An isolated polynucleotide comprising a nucleotide sequence which is a polymorphic variant of a reference sequence for the IGERB cDNA or a fragment thereof, wherein the reference sequence comprises SEQ ID NO:2 and the polymorphic variant comprises at least one polymorphism selected from the group consisting of thymine at a position corresponding to nucleotide 132, thymine at a position corresponding to nucleotide 428, guanine at a position corresponding to nucleotide 473, and guanine at a position corresponding to nucleotide 532 .

8. A recombinant organism transformed or transfected with the isolated polynucleotide of claim 7, wherein the organism expresses an immunoglobulin E receptor beta chain (IGERB) protein encoded by the polymorphic variant sequence.

9. The recombinant organism of claim 8 which is a nonhuman transgenic animal.

10. An isolated polypeptide comprising an amino acid sequence which is a polymorphic variant of a reference sequence for the IGERB protein or a fragment thereof, wherein the reference sequence comprises SEQ ID NO: 3 and the polymorphic variant comprises one or more variant amino acids selected from the group consisting of methionine at a position corresponding to amino acid position 143, cysteine at a position corresponding to amino acid position 158, and alanine at a

position corresponding to amino acid position 178 .

11. An isolated antibody specific for and immunoreactive with the isolated polypeptide of claim 10.
12. A method for screening for drugs targeting the isolated polypeptide of claim 10 which comprises contacting the IGERB polymorphic variant with a candidate agent and assaying for binding activity.
13. A composition comprising at least one genotyping oligonucleotide for detecting a polymorphism in the immunoglobulin E receptor beta chain (IGERB) gene at a polymorphic site selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.
14. The composition of claim 13, wherein the genotyping oligonucleotide is an allele-specific oligonucleotide that specifically hybridizes to an allele of the IGERB gene at a region containing the polymorphic site.
15. The composition of claim 14, wherein the allele-specific oligonucleotide comprises a nucleotide sequence selected from the group consisting of of SEQ ID NOS:4-25, the complements of SEQ ID NOS: 4-25, and SEQ ID NOS:26-69.
16. The composition of claim 13, wherein the genotyping oligonucleotide is a primer-extension oligonucleotide.
17. A method for genotyping the immunoglobulin E receptor beta chain (IGERB) gene of an individual, comprising determining for the two copies of the IGERB gene present in the individual the identity of the nucleotide pair at one or more polymorphic sites (PS) selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.
18. The method of claim 17, wherein the determining step comprises:
    (a) isolating from the individual a nucleic acid mixture comprising both copies of the IGERB gene, or a fragment thereof, that are present in the individual;
    (b) amplifying from the nucleic acid mixture a target region containing at least one of the polymorphic sites;
    (c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;
    (d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and
    (e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.
19. A method for haplotyping the immunoglobulin E receptor beta chain (IGERB) gene of an individual which comprises determining, for one copy of the IGERB gene present in the individual, the identity of the nucleotide at one or more polymorphic sites (PS) selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13.
20. The method of claim 19, wherein the determining step comprises

(a) isolating from the individual a nucleic acid molecule containing only one of the two copies of the IGERB gene, or a fragment thereof, that is present in the individual;

(b) amplifying from the nucleic acid molecule a target region containing at least one of the polymorphic sites;

(c) hybridizing a primer extension oligonucleotide to one allele of the amplified target region;

(d) performing a nucleic acid template-dependent, primer extension reaction on the hybridized genotyping oligonucleotide in the presence of at least two different terminators of the reaction, wherein said terminators are complementary to the alternative nucleotides present at the polymorphic site; and

(e) detecting the presence and identity of the terminator in the extended genotyping oligonucleotide.

21. A method for predicting a haplotype pair for the immunoglobulin E receptor beta chain (IGERB) gene of an individual comprising:

(a) identifying an IGERB genotype for the individual at two or more of polymorphic sites selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13;

(b) enumerating all possible haplotype pairs which are consistent with the genotype;

(c) accessing data containing the IGERB haplotype pairs determined in a reference population; and

(d) assigning a haplotype pair to the individual that is consistent with the data.

22. A method for identifying an association between a trait and at least one genotype or haplotype of the immunoglobulin E receptor beta chain gene which comprises comparing the frequency of the genotype or haplotype in a population exhibiting the trait with the frequency of the genotype or haplotype in a reference population, wherein the genotype or haplotype comprises a nucleotide pair or nucleotide located at one or more polymorphic sites selected from PS2, PS3, PS4, PS5, PS6, PS7, PS8, PS9, PS10, PS12, and PS13, wherein a higher frequency of the genotype or haplotype in the trait population than in the reference population indicates the trait is associated with the genotype or haplotype.

23 The method of claim 22, wherein the haplotype is selected from haplotype numbers 1-12 shown in Table 5.

24. The method of claim 23, wherein the trait is a clinical response to a drug targeting IGERB .

25. A computer system for storing and analyzing polymorphism data for the immunoglobulin E receptor beta chain gene, comprising:

(a) a central processing unit (CPU);

(b) a communication interface;

(c) a display device;

(d) an input device; and

(e) a database containing the polymorphism data;

wherein the polymorphism data comprises the genotypes and haplotype pairs shown in Table 4 and the haplotypes shown in Table 5.

26. A genome anthology for the immunoglobulin E receptor beta chain (IGERB) gene which comprises IGERB isogenes defined by haplotypes 1-12 shown in Table 5.

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
1 March 2001 (01.03.2001)

PCT

(10) International Publication Number
WO 01/14588 A1

(51) International Patent Classification[7]: C12Q 1/68, C12N 15/00, G01N 33/53, C07K 16/00, C07H 21/02, 21/04, A01K 67/00, 67/033

(21) International Application Number: PCT/US00/22175

(22) International Filing Date: 11 August 2000 (11.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
60/150,423 24 August 1999 (24.08.1999) US

(71) Applicants *(for all designated States except US)*: **GENAISSANCE PHARMACEUTICALS, INC.** [US/US]; Five Science Park, New Haven, CT 06511 (US). **NANDABALAN, Krishnan** [IN/US]; 228 Village Pond Road, Guilford, CT 06437 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **DENTON, R., Rex** [US/US]; 129 Hunters Trail, Madison, CT 06443 (US). **KLIEM, Stefanie, E.** [DE/US]; 1298 Hartford Turnpike, North Haven, CT 06473 (US). **STEPHENS, Joel, Claiborne** [US/US]; 46 Crabapple Lane, Guilford, CT 06437 (US).

(74) Agent: **HENDERSON, Melodie, W.**; Genaissance Pharmaceuticals, Inc., Five Science Park, New Haven, CT 06511 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *With international search report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/14588 A1

(54) Title: DRUG TARGET ISOGENES: POLYMORPHISMS IN THE IMMUNOGLOBULIN E RECEPTOR BETA CHAIN GENE

(57) Abstract: Polynucleotides comprising one or more of 11 novel single nucleotide polymorphisms in the human immunoglobulin E receptor beta chain (IGERB) gene are described. Compositions and methods for detecting one or more of these polymorphisms are also disclosed. In addition, various genotypes and haplotypes for IGERB gene that exist in the population are described.

POLYMORPHISMS IN THE IGERB GENE

```
AAGCTTTTCA AAGGTGCAAT TGGATAACTT CTGCCATGAG AAATGGCTGA
ATTGGGACAC AAGTGGGGAC AATTCCAGAA GAAGGGCACA TCTCTTTCTT      100
TTCTGCAGTT CTTTCTCACC TTCTCAACTC CTACTAAAAT GTCTCATTTT
CAGGTTCTGT AAATCCTGCT AGTCTCAGGC AAAATTATGC TCCAGGAGTC      200
TCAAATTTTC TTATTTCATA TTAGTCTTTA TTTAGTAGAC TTCTCAATTT
                                                      T
TTCTATTCAT CACAAGTAAA AGCCTGTTGA TCTTAATCAG CCAAGAAACT      300
                 A
TATCTGTCTG GCAAATGACT TATGTATAAA GAGAATCATC AATGTCATGA
GGTAACCCAT TTCAACTGCC TATTCAGAGC ATGCAGTAAG AGGAAATCCA      400
CCAAGTCTCA ATATAATAAT ATTCTTTATT CCTGGACAGC TCGGTTAATG
AAAAAATGGA CACAGAAAGT AATAGGAGAG CAAATCTTGC TCTCCCACAG      500
     [exon 1: 456..
GAGCCTTCCA GGTAGGTACA AGGTATTATT TTTTTCTACC CTCAGTCACT
                 ..511]
TGTGGCAGGG GAAGTCATAG TCACGGTGCT TAGGAGATGA AACTTTATTG      600
ATTTAGGCAT GGATCCATCT AGTTTAATTA ATATATTGGG TATGAGGAAG
CTACTTGCTG TACTTTCCAT GTGGTTCTCT CTCCCTGGAG AGGAACATTT      700
TTACTCAGCT TGCAAACTGG AAATAGATTT TCTCACATTA GAAGCTCATT
TTCTGGGTAT GAGACAGGAG AGTTCATACT GTGTATGTAG ATCTCTGGCT      800
TCTGGGTCTG ACATGTGCTG AGGGACACAT ATCCTTCACA CATGCTTTTA
TAAATACTTG ATAAAGTAAC CTGCTTCTTG ATTGGTCTTT ATAATCCATA      900
AGCTGTGGGA TGCTTCTCTG AAGATGAAAA TAGTAATAGA GTCCCATCTA
GCTATTCAAA GCCATTCCTT CATTGTATTC TGTGCACATG AAGTTGGGGT     1000
TTGTTACTGA CAAAATATAT TCAGATACAT TTCTATGTTA AAAGGATTGT
GAGATGCATA GGTAAATGTG TTTATTTTCA GTTTTACTTG TCAACATAGA     1100
TGAATGAGAA AGAACTTGAA AGTAACACTG GATTAAGAAT AGGAAAATTT
GGCATGGATT TTGCTCCATT TTGTCCCATC TAATCACTTG GATAGTGTTC     1200
AGGTGTTCTT GGTCAGTTAC TTGGATGCTC TGAGCTTTAG TTTCTTGGTG
ATTACAATGA AGATTTGAAT TACAGGATGG CTTTGAAAAA ATAAACAAAA     1300
CTCCCCTTTC TGTCTGTCGA GAATGTTGCA CAGGGAGTTA CAGAATGTTC
TCATGACTGA ATTGCTTTTA AATTTCACAG TGTGCCTGCA TTTGAAGTCT     1400
     [exon 2: 1381..
TGGAAATATC TCCCCAGGAA GTATCTTCAG GCAGACTATT GAAGTCGGCC
TCATCCCCAC CACTGCATAC ATGGCTGACA GTTTTGAAAA AAGAGCAGGA     1500
     T
GTTCCTGGGG GTGAGTGAGC CTCCTCCAAC TTTGACTAGA GTAAGGGTTG
                 ..1510]
GGTCTAGAAA AGAATATTGA GTTGCATCAA CTGTTTTCCC ACTTGGATTC     1600
ATGAGAGGTG TTAGGTCCTT TAAAAAACAT GGTAGATAAA GAGTTGACAC
TAACTGGGTC CTTTTGGGAA GAGCCAGAAG CATTTCCTCA TAAAGACTTT     1700
AAATTGCTAG GACGAGAATG GCCAACAGGA GTGAAGGATT CATAACTTTA
TCTTTACTTA GATGTAAAGA ACAATTACTG ATGTTCAACA TGACTACATA     1800
CATAAAGGCG CATGGAGAAA AGTATTGGCC TTCCATGCAT TAGGTAGTGC
TTGTATCAAT TCTTATAGTG GCTAGGGTAT CCTGGAAAAT CTTACGTGTG     1900
GATCATTTCT CAGGACAGTC TAGGACACTA ACGCAGTTTC TCATGTTTGG
CTTCTATTAT TAAAAAATGA TACAATCTCG GGAAAATTTT TTTGATTTTC     2000
ATGAAATTCA TGTGTTTTTC TATAGGTAAC ACAAATTCTG ACTGCTATGA
     [exon 3: 2026..
TATGCCTTTG TTTTGGAACA GTTGTCTGCT CTGTACTTGA TATTTCACAC     2100
ATTGAGGGAG ACATTTTTTC ATCATTTAAA GCAGGTTATC CATTCTGGGG
```

FIGURE 1A

```
AGCCATATTT GTGAGTATAT ATCTATAATT GTTTCTGAAA TAACACTGAA        2200
                ..2160]
CATAGGTTTT TCTCTTTCTC AGATCTAACC AGTTGTTTAT TCCCAGTATT
AAGATGATAT TTATAATTCT TAATTATAAA TATATGTGAG CATATATAAC        2300
  C
ATAGATATGC TCATTAACAA CAACAAAAGA TTCTTTTTAC AATTAACGGT
 C
GGGTTAAACA TTTAGCCCAC AGTTTTATCC CATGAGAAAC CTGAATCTAA        2400
TACAAGTTAA ATGACTTGCC TAAGGGCCAC TTGACTAATA GTAATTGAAC
CTAAACTTTC AGAATCCAAC TCCAGGAACA TACTTCTAGC ACTATTCATC        2500
AATAAAGTTA TATGATAAAT ACATACAACT TTATCTGTCA ACTAAAAATA
ACAACAGAGG CTGGGCATGG TGGCTCACAC CCGTAATCCC AGCACTTTGG        2600
GAGGCTGAGG CAGGTGGATC ACCTGAGGTC AGGAGTTTGA GACCAGCCTG
ACCAACATGG TGAAACCTCA TCTCTACTAA ATATAAAAAA TTAGCTGAGT        2700
GTGATAGTGC ATACCTGTAA TCCAGCTACT TAAGAGGCTG AGGCAGGAGG
CTTGTTTGAA CCTGGAAGGC AGAGGTTGCA GTGAGCTGAG ATTGTGGCAT        2800
TGCACTCCAG CCTGGGCAAT AAGTGCGAAC TCTGTCTCAA AATAATAATA
ATAATAATAG AAAATAAAGT TGTCTTCATG AAAAATGAGG AAAGAGATTG        2900
CTGGGGTGAG AAACATTAAG ATCAATGGGC ATATGGTGAC CTTCTATGCC
CTAGAAACTC TTTTANGGTA TTTTCTCCTG GTATCTCTTT TACNCATCGT        3000
TCTATCTGGA AAAATAGGTG GATGAGTGAG ATAATAACGG TATATACTTT
TTAAAGGTCT AATTGACATA TATAAATTGC AAGTATTTCA GATGTCAATT        3100
TGCTAACCTT GACACACATA GACACACATG AAAACATCAC CACATTAATA
CAATGTATGT ATCCATCATT CCAAAAGCTT CCCTGTGTAT CTTTGTAACT        3200
CTTTCTTCCT CCCTCCACTC CTTGTCCTCT CGTTCCCAAG AAAACATTGA
TCTGCTTCCT GTGAATATAA ATTAACTTAC ATTTTTTAGA GCTTTATATA        3300
AGTATGTTCT CTTTACTGTT TGTCTTCCTT CGCTGCACAG TTATTTTGAG
ATTCTTCAAG TTTTTTCTTT ATATCGATAC TTCATTCACA AGAATATATT        3400
TTAATTCTAG ACTATGTCAC ATTGACTTTG TCGTCTGCTA AATCCTTAGT
GCTCAGATGA CTTGTTCAGG ACTCTCCTTG AACCTGTACC TCTGTTANAT        3500
TGAAACTTGT CTCTACTGTC TTTTTATTTC AAACACAGCT TATTAGGTGT
CTCTCAACCC ATCAAACNCA CAATCTGAGT CTTTAGGAGA TTGCTTTGAA        3600
TTTGTGCTAT TGACTTATAT NTATATNAAA TNTGTAAATG TTTGGTAAAA
ATATCATCAT GTACNTTTTC ATAATTACGC TATNTNCACA TGATATATGT        3700
CAGACTCTGG AAATATGCAT GCCACAGACA CGTGTTTCTT GCCTAAAGGG
GCTGATGGAA GACNCACATA CNAATAGACG ATTGCAGTAG AATGAGAGTG        3800
GTGGTCTAAN CAGTACATGT CCTGATGTTG CTCGGACAGT TACTACNCCA
AGAGTACCCC CTGCATTGTC AGGGTTAGCA TCTCCTGGAA GCCTCATGTA        3900
AATGAAGAAT TTCATGCTCC ATCCAGGACC TAATGAATAA GAATCTGCAT
TTTAGCAAGA CCCTCATATG ATTCATATAC ACTTTTTTTT TTTTTTTTTA        4000
GATGGAGTCT CACTCTTGTC GCCCAGGCTG GAGTGCAATG GCATGATCTT
GGCTCACTGC AACCTCTGCC TCCCGGGTTC AAGTGATTCT CCTGTCTCAG        4100
CCTCCCTAGT AGCTGGGACT ACAGGTGCAT GCCACAGTGG CTGGCTAATT
TTTGTATTTT TAGTAGAGAC AGGGTTTCAC CATTTTGGTC AGGCTGGTCT        4200
TGAACTCATG ACCTCCGGTG ATTCCCCCGC CTCGGCTTCC CAAAGTGCTG
GGATTACAGA CATGAGCCAC CACACCCGCC TTATTCGTAT ACNCATTTAA        4300
TTCTGAGAAG CACTCTATAG AAAATAAGAA TAAGAAAATA TTGGGCTCAC
AGGTGACATT AATAAGTAAC TTTATCGAGT ACCCCAAATT TTACCTATGT        4400
TTGGAAGATG GGGTTAAAAG GACACATTGA AAACAAGAAC TCATTGTGGC
TTTTTTTTCC TCCTTTTTGA ACAGTTTTCT ATTTCTGGAA TGTTGTCAAT        4500
     [exon 4: 4475..
TATATCTGAA AGGAGAAATG CAACATATCT GGTGAGTTGC CCGTTTCTGT
                ..4531]
```

FIGURE 1B

```
CTTTGTCCAT CCTTGAAAAG ATAAGAAGAA CAGAGTTTTA AGAGTCTTAA      4600
GGGAAACACA TCTTTGTCTC CTATATTACT TGTGAATGTG GATATATGAT
TTTGTTTCAA TCTATTTTGT GTCCTAAGGC TTTTTGCAAC AGAAGTTGGA      4700
TATATCATTA GAAACATAAA TTGTACCATT TAACATACAT GAAGTTTATG
TTTACCTTGA CGTTCTTCTA AAAAGTGTCC TACACCGGCA TTGTCCTTGT      4800
AGGCATATTC ACATGATCAA ATAAAATAAT TAGTTTTCAA TTAAGGAGAA
TATTTGAGGA AAGACCGTAC GTGTTCATGT GGTTCCTGAA GGCAGTCCAG      4900
TGAGAAAGTA ATATATGCTT CATTAAACAA TGCGGACATT TTCAGGGTTT
CCCTTTTTAA CCAAAATTTG GAAGCAATGT GGAATTTACT GGATGCATCC      5000
AGCCCTGAAA TGAAGATAGG TTTATTGAAT GTGCCAGCAA GTGCAGGCCC
AGGTCTGAGT GTTCTTCATT ATTATCAGGT GAGAGGAAGC CTGGGAGCAA      5100
     [exon 5: 5079..
ACACTGCCAG CAGCATAGCT GGGGGAACGG GAATTACCAT CCTGATCATC
                              T
AACCTGAAGA AGAGCTTGGC CTATATCCAC ATCCACAGTT GCCAGAAATT      5200
                        G
TTTTGAGACC AAGTGCTTTA TGGCTTCCTT TTCCACTGTA TGTATTTTTT
                                 G
                ..5237]
TTTGTGTGGG AAGACTAAGA TTCTGGGTCC TAATGTAAGT AAGAAGCCCT      5300
 G   A
CTTCTCCTGT TCCATGAACA CCATCCTTTT CTGTAACTTC TATTACACAG
TATAGTGGTT CTGTAAGTTC ACACAGCCCA GGGAGATGCT GGCTGCCCAC      5400
TCCCCTCAAC CCAGGCAAAT TCCTCGGGGT TAAAGTTATC TACTGCAAGT
GACGATCTCT GGGTTTTTCT GTGCCTGTGT TTGTGTGTGT GTGTGTGTGT      5500
GTGTGTGTGT GTATGTGTCA CTTTAAAAGG ACTGGTCAGA TGGTAGGGAG
ATGAAAACAG GAGATGCTAT AAGAAAATAA ACTTTTGGGG CGAATACCAA      5600
TGTGACTCTT TTTGTTTGTC ATTTGTTGCT GTTCAATAGG AAATTGTAGT
     [exon 6: 5640..
GATGATGCTG TTTCTCACCA TTCTGGGACT TGGTAGTGCT GTGTCACTCA      5700
CAATCTGTGG AGCTGGGGAA GAACTCAAAG GAAACAAGGT AGATAGAAGC
                ..5738]
CCGATATAAA ATCTTGAATG ACAGGTTAAC GAATTGGAGC TTTATTCCTT      5800
AAAATATGGC CTGGGTTTTC TGAAACATTT CTTCCAGAAA ATAGTTTCTC
CAAGTTTTAT TACTTTGGTT TACAAATCTC ACATTTAAAT CACATTTTAT      5900
ACCATAAGTA GCACACATTT CATAATATTC CTCTGAATGA GGGTTGGGAT
AATAGGACTG ATATGTTAGA AATGCCTTAA AGTGTGTGGA GCATGAGAGA      6000
TGGATGTACA GAAGGCTTGT GAGGAAACCA CCCAGGTATC TGGCCTTGTT
TTCTGCCCCA GAACTAGCCG CCTATTCCTG TTTCTGTTTT ATTCCTTTGT      6100
TTCTTGACTT TTCCTTTCCA ACTTGCTCTA AAACCTCAGT TTTCTTTCCT
TTCTGATTCA TGACTACCAA ATGTTTTCAC TTGCCTCACC CGTCCATTAC      6200
ACCTTTGATA AGAACCACCA GACCTTGTGC TCATGTACTT GCCCATGTCT
GATGGAAGAA ACATACTCTC TCCATCTGTC CACTTTCCTG AGGCATTCAA      6300
GTCTAGCCAC CTTTTAAAAT CACTCTCCTC CAGGCTGGGC ACGGTGTCAC
GCCTGTAATC TCAGCACTTT GTGAGGCTGA GGAGGGCGGA TCACTTGAAG      6400
TCAGGAGTTC AAAACCAGCC TGGCCAAATG GCAAAACCAA ATCTTCTTCA
ATTATAACCA AATCTTAAAC CAAATCTCTA CTAAAAAATA CAACAAAACA      6500
AAACAACAAC AACAAAAACA GAAAAGGAAA CATTAGCCCA GCGTGGTGGC
AGGTACCTGA GGTTCCAGAT ACTTGGGAGG CTGAAGCAGG AGAATCGCTT      6600
GAGCCCAAGA GATGGAGGTT GCAGTGAGCC GAGATCATGC CACTGCACCA
CAGCCAGGGT GACAGAGCCA TACTTCCCAG CACATTGGGA GGCCAAAGCT      6700
GAAGAATAAT TTGAGGTGAG GATTTGGAGA CCAGCCTGGC CAACATGGTG
AAACTCCGTC TGTACTAAAA ATATAAAACT TAGTGGGGCA TGGGGGCACA      6800
```

FIGURE 1C

```
CACCTGTAAT TTCAGCTACT TAGGAGGCTG AGGCAGGAGA ATTGCTTGAA
CCCGGGAGGC GGAAGTTGCA GTGAGCCAAG ATCGTGGCCA CTGCACTCCA      6900
GCCTGGGTGA CATAGTGAGA TTCTGTCTCA AAAAAAATAA AAGAAATTTA
AAAAATCACT CTCTTCCAAA GATAGATAAA TAAGACAGCA GATATACTAA      7000
GGAATAACCT CACCAACTTG TCATTGACTG ACATGATTTC TTTTGGCCCA
CTTGGCCAGC TAGTCTGGTT TGGTTTTCTG GAAATGAAAG AAATAATCAG      7100
AGTTTAATGA CAGAGAGCGT GAGACCCAGA AAGACAAAAG TAGATGAGGT
AAGTCTCTTG AGCGAGACTT CTAGGGATGG GAAATTTGTG GTGATTGATA      7200
TGAAATGATT TTTCCCTTAT CAGGTTCCAG AGGATCGTGT TTATGAAGAA
     [exon 7: 7224..
TTAAACATAT ATTCAGCTAC TTACAGTGAG TTGGAAGACC CAGGGGAAAT      7300
                                                    G
GTCTCCTCCC ATTGATTTAT AAGAATCACG TGTCCAGAAC ACTCTGATTC
                 ..7322]
ACAGCCAAGG ATCCAGAAGG CCAAGGTTTT GTTAAGGGGC TACTGGAAAA      7400
ATTTCTATTC TCTCCACAGC CTGCTGGTTT TACATTAGAT TTATTCGCCT
C                                                     T
GATAAGAATA TTTTGTTTCT GCTGCTTCTG TCCACCTTAA TATGCTCCTT      7500
CTATTTGTAG ATATGATAGA CTCCTATTTT TCTTGTTTTA TATTATGACC
ACACACATCT CTGCTGGAAA GTCAACATGT AGTAAGCAAG ATTTAACTGT      7600
TTGATTATAA CTGTGCAAAT ACAGAAAAAA AGAAGGCTGG CTGAAAGTTG
AGTTAAACTT TGACAGTTTG ATAATATTTG GTTCTTAGGG TTTTTTTTTT      7700
TTTTAGCATT CTTAATAGTT ACAGTTGGGC ATGATTTGTA CCATCCACCC
ATACCCACAC AGTCACAGTC ACACACACAT ATGTATTACT TACACTATAT      7800
ATAACTTCCT ATGCAAATAT TTTACCACCA GTCAATAATA CATTTTTGCC
AAGACATGAA GTTTTATAAA GATCTGTATA ATTGCCTGAA TCACCAGCAC      7900
ATTCACTGAC ATGATATTAT TTGCAGATTG ACAAGTAGGA AGTGGGGAAC
TTTTATTAAG TTACTCGTTG TCTGGGGAGG TAAATAGGTT AAAAACAGGG      8000
AAATTATAAG TGCAGAGATT AACATTTCAC AAATGTTTAG TGAAACATTT
GTGAAAAAAG AAGACTAAAT TAAGACCTGA GCTGAAATAA AGTGACGTGG      8100
AAATGGAAAT AATGGTTATA TCTAAAACAT GTAGAAAAAG AGTAACTGGT
AGATTTTGTT AACAAATTAA AGAATAAAGT TAGACAAGCA ACTGGTTGAC      8200
TAATACATTA AGCGTTTGAG TCTAAGATGA AAGGAGAACA CTGGTTATGT
TGATAGAATG ATAAAAAGGG TCGGGCGCGG AGGCTCACGC CTGTAATCCC      8300
AGCCCTTTGG GAGGCCGAGG TGGGCAGATC ACGAAGTCAG TAGTTTGAGA
CCAGCCTGGC CAACATAGTG AAACCCCGTC TCTACTAAAA ATACAAAAAA      8400
AAAATTAGCT GGGTGTGGTG GCAGTCACCT GTAGTCCCAG CTACTTGGGA
GGATGAGGCA GGAGAATCGC TTGAACCTGG GAGGCGGAGG TTGCAGTGAG      8500
CCGAGATCGC ACCAGTGCAC TCCAGCCTTG GTGACAATGG GAGACTCCAT
CTCAAAAAAA AAAAAAAAAA AAAAAAGATA AAAAGTCAGA AATCTGAAAA      8600
GTGGAGGAAG AGTACAAATA GACCTAAATT AAGTCTCATT TTTTGGCTTT
GATTTTGGGG AGACAAAGGG AAATGCAGCC ATAGAGGGCC TGATGACATC      8700
CAATACATGA GTTCTGGTAA AGATAAAATT TGATACACGG TTTGGTGTCA
TTATAAGAGA AATCATTATT AAATGAAGCA AGTTAACACT CTAAGAGAAT      8800
TATTTTGAGA TAGAAGTGAA GCTAAGCTAA ACTTCACATG CCTATAATTG
GAGGGAAAAA CTAAGGATAA AATCTAGCCT AGAAGATACA ATAATTAGTC      8900
ATAAACATGC ATTGTGAAAC TGTAGAGAGC AGGTAGCCCA AAATAGAGAA
AGATTAGATA AAGAGAAAAT AAGTATCCAT CAGAGACAGT ATCTCTAGGC      9000
TTGGGCAAGA GAAAAGTCCA CAGTGATAAG CAACTCCACC TAAGGCATGA
ATATGCGGCA GAGAAAACAG CAATAGTGAA TGAATGCAAA AGGTGCTGAG      9100
CAAATTCCAC ACATGAGTAT TGTGCATGAG TAAATGAATA AAACATTTGC
AAAGACCTTT AGAGAAAGAG AATGGGAGCA TATGTGCGAA ATAAGATAGT      9200
TGATTATGAA TAGAAGGTAG TGAAGAAAAG CAAGCTAAGA AAAAATTCTG
```

FIGURE 1D

```
TTTATAAAAG AAGGAAAAGA TAGTTTATGT TTTTAGCCTA AGTATAAGAG      9300
TCCTACAGAT GGACTGAAAA AAATCAGTCT GAGAGTATTA GTCACAATTA
ATGAAATAAT TACATTTTAT GTATTGAGGA TGCCAAGATT AAAAGGTGAC      9400
AGGTAGATGT TAATTTCCCT AGATTGTGAA AGTGATCACG ACAATCACAC
AACAAATAAT TAAGTGACTT GGTATGCTTT ATTTAATTGT AGGGCCTGAG      9500
GTTTTCCATT CTCATTTTTC TAAAATACAA TTTTGTTTCT CCAAATTTGA
CAGCAGAATA AAAACCCTAC CCTTTCACTG TGTATCATGC TAAGCTGCAT      9600
CTCTACTCTT GATCATCTGT AGGTATTAAT CACATCACTT CCATGGCATG
GATGTTCACA TACAGACTCT TAACCCTGGT TTACCAGGAC CTCTAGGAGT      9700
GGATCCAATC TATATCTTTA CAGTTGTATA GTATATGATA TCTCTTTTAT
TTCACTCAAT TTATATTTTC ATCATTGACT ACATATTTCT TATACACAAC      9800
ACACAATTTA TGAATTTTTT CTCAAGATCA TTCTGAGAGT TGCCCCACCC
TACCTGCCTT TTATAGTACG CCCACCTCAG GCAGACACAG AGCACAATGC      9900
TGGGGTTCTC TTCACACTAT CACTGCCCCA AATTGTCTTT CTAAATTTCA
ACTTCAATGT CATCTTCTCC ATGAAGACCA CTGAATGAAC ACCTTTTCAT     10000
CCAGCCTTAA TTTCTTGCTC CATAACTACT CTATCCCACG ATGCAGTATT
GTATCATTAA TTATTAGTGT GCTTGTGACC TCCTTATGTA TTCTCAATTA     10100
CCTGTATTTG TGCAATAAAT TGGAATAATG TAACTTGATT TCTTATCTGT
GTTTGTGTTG GCATGCAAGA TTTAGGTACT TATCAAGATA ATGGGGAATT     10200
AAGGCATCAA TAAAATGATG CCAAAGACCA AGAGCAGTTT CTGAAGTCCT
CCTTTTCATC AGCTCTTTAT CAAACAGAAC ACTCTATAAA CAACCCATAG     10300
CCAGAAAACA GGATGTAGGA ACAATCACCA GCACACTCTA TAAACAACCC
ATAGCCAGAA AACAGAATGT AAGGACAATC ACCAGCCATC TTTTGTCAAT     10400
AATTGATGGA ATAGAGTTGA AAGGAACTGG AGCATGAGTC ATATTTGACC
AGTCAGTCCT CACTCTTATT TACTTGCTAT GTAAACTTGA GAAAGCTTTT     10500
TTCTCTTTGT GAACCTCAGG TTTTACATCT GAAAATGAGA AATTTGGAAC
AAAAGATTCC TAACTGGTCT TTCTGTTCCC ATATTCTGTG ATTTTTCAAT     10600
ATTTAGGATT TTTGGTAATC ACAATTACTT AGTTTGTGGT TGAGATAGCA
ACACGAATCA GAACTATTTG GTGGACATAT TTTCAAAGGA GTAGCTCTCC     10700
ACTTTGGGTA AAGAAGTGAT GCNGGTCGTG GTGGCTCACG CCTGTAATCC
CAGCACTTTA GGGAGGCCAA GGCGGGTGGA TCACGAGGTC AGGAGATCGA     10800
GACCATCCTG GCTAACACGG TGAAACCCCG TCTCTACTAA AAAATACAAA
AAATTAGCCA GGCGTGGTGG CGGGCGCCTG TAGTCCCACG TACTCGGGAG     10900
GCTGAGGCAG GAGAATGGCA TGAACCAGGG AGGCGGAGCT TGCCGTGAGC
CGAGATAGCG CCACTGCAGT CCCTCCTGGG CAAAAGAGCA AGACTGCGTC     11000
TCAAAAAAAA AAAAAAAAAA AAAAAAAGAA GTGTGTGGAG TAGCAGGACA
CCTGCAACAA TAATATTTTT CTAAATCCCT CTGAAAAATG CTAATCAAAG     11100
GGTTTTTTTC CTAAAAATTG TCTTAGAAAT AAAATTTCCC CTTTGGGAGA
CCGAGGCTGG CAGATCACGA GGTCAGGAGA TAGAGACCAC GGTGAAACCC     11200
CGTCTCTACT AAAAATACTA AAAATTAGCC GGGGNGTGGT GGTGGGTACA
CCTGTAGTCC CAGCTACTTG GAGGCTGAGG CTGGAGAATC ACGTGAAC       11298
```

FIGURE 1E

POLYMORPHISMS IN THE CODING SEQUENCE OF IGERB

```
ATGGACACAG AAAGTAATAG GAGAGCAAAT CTTGCTCTCC CACAGGAGCC
TTCCAGTGTG CCTGCATTTG AAGTCTTGGA AATATCTCCC CAGGAAGTAT      100
CTTCAGGCAG ACTATTGAAG TCGGCCTCAT CCCCACCACT GCATACATGG
                                    T
CTGACAGTTT TGAAAAAAGA GCAGGAGTTC CTGGGGGTAA CACAAATTCT      200
GACTGCTATG ATATGCCTTT GTTTTGGAAC AGTTGTCTGC TCTGTACTTG
ATATTTCACA CATTGAGGGA GACATTTTTT CATCATTTAA AGCAGGTTAT      300
CCATTCTGGG GAGCCATATT TTTTTCTATT TCTGGAATGT TGTCAATTAT
ATCTGAAAGG AGAAATGCAA CATATCTGGT GAGAGGAAGC CTGGGAGCAA      400
ACACTGCCAG CAGCATAGCT GGGGGAACGG GAATTACCAT CCTGATCATC
                              T
AACCTGAAGA AGAGCTTGGC CTATATCCAC ATCCACAGTT GCCAGAAATT      500
                   G
TTTTGAGACC AAGTGCTTTA TGGCTTCCTT TTCCACTGAA ATTGTAGTGA
                                    G
TGATGCTGTT TCTCACCATT CTGGGACTTG GTAGTGCTGT GTCACTCACA      600
ATCTGTGGAG CTGGGGAAGA ACTCAAAGGA AACAAGGTTC CAGAGGATCG
TGTTTATGAA GAATTAAACA TATATTCAGC TACTTACAGT GAGTTGGAAG      700
ACCCAGGGGA AATGTCTCCT CCCATTGATT TATAA                      735
         G
```

FIGURE 2

ISOFORMS OF THE IGERB PROTEIN

```
MDTESNRRAN LALPQEPSSV PAFEVLEISP QEVSSGRLLK SASSPPLHTW
LTVLKKEQEF LGVTQILTAM ICLCFGTVVC SVLDISHIEG DIFSSFKAGY      100
PFWGAIFFSI SGMLSIISER RNATYLVRGS LGANTASSIA GGTGITILII
                                            M
NLKKSLAYIH IHSCQKFFET KCFMASFSTE IVVMMLFLTI LGLGSAVSLT      200
      C                      A
ICGAGEELKG NKVPEDRVYE ELNIYSATYS ELEDPGEMSP PIDL            244
                                   G
```

FIGURE 3

Practitioner's Docket No. MWH-008US *PATENT*

COMBINED DECLARATION AND POWER OF ATTORNEY

(ORIGINAL, DESIGN, NATIONAL STAGE OF PCT, SUPPLEMENTAL, DIVISIONAL, CONTINUATION, OR C-I-P)

As a below named inventor, I hereby declare that:

TYPE OF DECLARATION

This declaration is for a national stage of PCT application.

INVENTORSHIP IDENTIFICATION

My residence, post office address and citizenship are as stated below, next to my name. I believe that I am an original, first and joint inventor of the subject matter that is claimed, and for which a patent is sought on the invention entitled:

TITLE OF INVENTION

Drug Target Isogenes: Polymorphisms in the Immunoglobulin E Receptor Beta Chain Gene

SPECIFICATION IDENTIFICATION

The specification is attached hereto:

AMENDMENT IDENTIFICATION

The PCT Article 34(2)(b) amendment filed with the International Bureau on March 26, 2001 is attached.

ACKNOWLEDGMENT OF REVIEW OF PAPERS AND DUTY OF CANDOR

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information, which is material to patentability as defined in 37, Code of Federal Regulations, Section 1.56, and which is material to the examination of this application, namely, information where there is a substantial likelihood that a reasonable Examiner would consider it important in deciding whether to allow the application to issue as a patent, and in compliance with this duty, there is attached an information disclosure statement, in accordance with 37 C.F.R. Section 1.98.

PRIORITY CLAIM (35 U.S.C. Section 119(a)-(d))

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s)

[redacted]

designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed.

Such applications have been filed as follows.

CLAIM FOR BENEFIT OF PRIOR U.S. PROVISIONAL APPLICATION(S)
(35 U.S.C. Section 119(e))

I hereby claim the benefit under Title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below:

PROVISIONAL APPLICATION NUMBER	**FILING DATE**
60/154,423	August 24, 1999

POWER OF ATTORNEY

I hereby appoint the following practitioner(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

APPOINTED PRACTITIONER(S)	REGISTRATION NUMBER(S)
Melodie W. Henderson	37,848
Inna Shtivelband	44,337
Gisela M. Field	47,562
Sandra L. Shaner	47,934

4

I hereby appoint the practitioner(s) associated with the Customer Number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

[redacted]

SEND CORRESPONDENCE TO
Inna Shtivelband
Genaissance Pharmaceuticals, Inc.
Five Science Park
New Haven, CT 06511

DIRECT TELEPHONE CALLS TO:
Inna Shtivelband
(203) 786-3529

DECLARATION

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

SIGNATURE(S)

R. Rex Denton
Inventor's signature _____
Date 7/18/2001 **Country of Citizenship** USA
Residence 129 Hunters Trail, Madison, CT 06443 CT

[redacted]

Stefanie E. Kliem
Inventor's signature _____
Date _____ **Country of Citizenship** Germany
Residence Kiefernweg 37, 61440 Oberusel, Germany

[redacted]

Krishnan Nandabalan
Inventor's signature _____
Date 7/19/2001 **Country of Citizenship** ~~India~~ USA
Residence 228 Village Pond Rd., Guilford, CT 06437 CT

[redacted]

J. Claiborne Stephens
Inventor's signature _____
Date 7-17-01 **Country of Citizenship** USA
Residence 46 Crabapple Lane, Guilford, CT 06437 CT

[redacted]

Practitioner's Docket No. MWH-008US *PATENT*

COMBINED DECLARATION AND POWER OF ATTORNEY

(ORIGINAL, DESIGN, NATIONAL STAGE OF PCT, SUPPLEMENTAL, DIVISIONAL, CONTINUATION, OR C-I-P)

As a below named inventor, I hereby declare that:

TYPE OF DECLARATION

This declaration is for a national stage of PCT application.

INVENTORSHIP IDENTIFICATION

My residence, post office address and citizenship are as stated below, next to my name. I believe that I am an original, first and joint inventor of the subject matter that is claimed, and for which a patent is sought on the invention entitled:

TITLE OF INVENTION

Drug Target Isogenes: Polymorphisms in the Immunoglobulin E Receptor Beta Chain Gene

SPECIFICATION IDENTIFICATION

The specification is attached hereto.

AMENDMENT IDENTIFICATION

The PCT Article 34(2)(b) amendment filed with the International Bureau on March 26, 2001 is attached.

ACKNOWLEDGMENT OF REVIEW OF PAPERS AND DUTY OF CANDOR

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose information, which is material to patentability as defined in 37, Code of Federal Regulations, Section 1.56, and which is material to the examination of this application, namely, information where there is a substantial likelihood that a reasonable Examiner would consider it important in deciding whether to allow the application to issue as a patent, and in compliance with this duty, there is attached an information disclosure statement, in accordance with 37 C.F.R. Section 1.98.

PRIORITY CLAIM (35 U.S.C. Section 119(a)-(d))

I hereby claim foreign priority benefits under Title 35, United States Code, Section 119(a)-(d) of any foreign application(s) for patent or inventor's certificate or of any PCT international application(s)

designating at least one country other than the United States of America listed below and have also identified below any foreign application(s) for patent or inventor's certificate or any PCT international application(s) designating at least one country other than the United States of America filed by me on the same subject matter having a filing date before that of the application(s) of which priority is claimed.

Such applications have been filed as follows.

CLAIM FOR BENEFIT OF PRIOR U.S. PROVISIONAL APPLICATION(S)
(35 U.S.C. Section 119(e))

I hereby claim the benefit under Title 35, United States Code, Section 119(e) of any United States provisional application(s) listed below:

PROVISIONAL APPLICATION NUMBER	**FILING DATE**
60/154,423	August 24, 1999

POWER OF ATTORNEY

I hereby appoint the following practitioner(s) to prosecute this application and transact all business in the Patent and Trademark Office connected therewith.

APPOINTED PRACTITIONER(S)	REGISTRATION NUMBER(S)
Melodie W. Henderson	37,848
Inna Shtivelband	44,337
Gisela M. Field	47,562
Sandra L. Shaner	47,934

I hereby appoint the practitioner(s) associated with the Customer Number provided below to prosecute this application and to transact all business in the Patent and Trademark Office connected therewith.

SEND CORRESPONDENCE TO
Inna Shtivelband
Genaissance Pharmaceuticals, Inc.
Five Science Park
New Haven, CT 06511

DIRECT TELEPHONE CALLS TO:
Inna Shtivelband
(203) 786-3529

DECLARATION

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

SIGNATURE(S)

R. Rex Denton
Inventor's signature _____
Date _____ **Country of Citizenship** USA
Residence 129 Hunters Trail, Madison, CT 06443

■■■■■■

Stefanie E. Kliem
Inventor's signature Stefanie Kliem
Date 12 July 2001 **Country of Citizenship** Germany
Residence Kiefernweg 37, 61440 Oberusel, Germany

■■■■■■

Krishnan Nandabalan
Inventor's signature _____
Date _____ **Country of Citizenship** India
Residence 228 Village Pond Rd., Guilford, CT 06437

■■■■■■

J. Claiborne Stephens
Inventor's signature _____
Date _____ **Country of Citizenship** USA
Residence 46 Crabapple Lane, Guilford, CT 06437

[redacted]

MWH0008
SEQUENCE LISTING

<110> Denton, R. Rex
Kliem, Stefanie E.
Nandabalan, Krishnan
Stephens, J. Claiborne

<120> DRUG TARGET ISOGENES: POLYMORPHISMS IN THE IMMUNOGLOBULIN E RECEPTOR BETA CHAIN GENE

<130> MWH-0008US IGERB
USS09889866

<140> TBA
<141> 2001-07-20

<150> 60/150,423
<151> 1999-08-24

<150> PCT/US00/22175
<151> 2000-08-11

<160> 145

<170> PatentIn Ver. 2.1

<210> 1
<211> 11298
<212> DNA
<213> Homo sapiens

<400> 1
aagcttttca aaggtgcaat tggataactt ctgccatgag aaatggctga attgggacac 60
aagtggggac aattccagaa gaagggcaca tctctttctt ttctgcagtt ctttctcacc 120
ttctcaactc ctactaaaat gtctcatttt caggttctgt aaatcctgct agtctcaggc 180
aaaattatgc tccaggagtc tcaaattttc ttatttcata ttagtcttta tttagtagac 240
ttctcaattt ttctattcat cacaagtaaa agcctgttga tcttaatcag ccaagaaact 300
tatctgtctg gcaaatgact tatgtataaa gagaatcatc aatgtcatga ggtaacccat 360
ttcaactgcc tattcagagc atgcagtaag aggaaatcca ccaagtctca atataataat 420
attctttatt cctggacagc tcggttaatg aaaaaatgga cacagaaagt aataggagag 480
caaatcttgc tctcccacag gagccttcca ggtaggtaca aggtattatt tttttctacc 540
ctcagtcact tgtggcaggg gaagtcatag tcacggtgct taggagatga aactttattg 600
atttaggcat ggatccatct agtttaatta atatattggg tatgaggaag ctacttgctg 660
tactttccat gtggttctct ctccctggag aggaacattt ttactcagct tgcaaactgg 720
aaatagattt tctcacatta gaagctcatt ttctgggtat gagacaggag agttcatact 780
gtgtatgtag atctctggct tctgggtctg acatgtgctg agggacacat atccttcaca 840
catgctttta taaatacttg ataaagtaac ctgcttcttg attggtcttt ataatccata 900
agctgtggga tgcttctctg aagatgaaaa tagtaataga gtcccatcta gctattcaaa 960
gccattcctt cattgtattc tgtgcacatg aagttggggt ttgttactga caaaatatat 1020
tcagatacat ttctatgtta aaaggattgt gagatgcata ggtaaatgtg tttattttca 1080
gttttacttg tcaacataga tgaatgagaa agaacttgaa agtaacactg gattaagaat 1140
aggaaaattt ggcatggatt ttgctccatt ttgtcccatc taatcacttg gatagtgttc 1200
aggtgttctt ggtcagttac ttggatgctc tgagctttag tttcttggtg attacaatga 1260
agatttgaat tacaggatgg ctttgaaaaa ataaacaaaa ctccccttcc tgtctgtcga 1320
gaatgttgca cagggagtta cagaatgttc tcatgactga attgctttta aatttcacag 1380
tgtgcctgca tttgaagtct tggaaatatc tccccaggaa gtatcttcag gcagactatt 1440
gaagtcggcc tcatccccac cactgcatac atggctgaca gttttgaaaa aagagcagga 1500
gttcctgggg gtgagtgagc ctcctccaac tttgactaga gtaagggttg ggtctagaaa 1560
agaatattga gttgcatcaa ctgttttccc acttggattc atgagaggtg ttaggtcctt 1620
taaaaaacat ggtagataaa gagttgacac taactgggtc cttttgggaa gagccagaag 1680
catttcctca taaagacttt aaattgctag gacgagaatg gccaacagga gtgaaggatt 1740
cataacttta tctttactta gatgtaaaga acaattactg atgttcaaca tgactacata 1800
cataaaggcg catggagaaa agtattggcc ttccatgcat taggtagtgc ttgtatcaat 1860

tcttatagtg gctagggtat cctggaaaat cttacgtgtg gatcatttct caggacagtc 1920
taggacacta acgcagtttc tcatgtttgg cttctattat taaaaaatga tacaatctcg 1980
ggaaaatttt tttgattttc atgaaattca tgtgtttttc tataggtaac acaaattctg 2040
actgctatga tatgcctttg ttttggaaca gttgtctgct ctgtacttga tatttcacac 2100
attgagggag acatttttttc atcatttaaa gcaggttatc cattctgggg agccatattt 2160
gtgagtatat atctataatt gtttctgaaa taacactgaa cataggtttt tctctttctc 2220
agatctaacc agttgtttat tcccagtatt aagatgatat ttataattct taattataaa 2280
tatatgtgag catatataac atagatatgc tcattaacaa caacaaaaga ttctttttac 2340
aattaacggt gggttaaaca tttagcccac agttttatcc catgagaaac ctgaatctaa 2400
tacaagttaa atgacttgcc taagggccac ttgactaata gtaattgaac ctaaactttc 2460
agaatccaac tccaggaaca tacttctagc actattcatc aataaagtta tatgataaat 2520
acatacaact ttatctgtca actaaaaata acaacagagg ctgggcatgg tggctcacac 2580
ccgtaatccc agcactttgg gaggctgagg caggtggatc acctgaggtc aggagtttga 2640
gaccagcctg accaacatgg tgaaacctca tctctactaa atataaaaaa ttagctgagt 2700
gtgatagtgc atacctgtaa tccagctact taagaggctg aggcaggagg cttgtttgaa 2760
cctggaaggc agaggttgca gtgagctgag attgtgccat tgcactccag cctgggcaat 2820
aagtgcgaac tctgtctcaa aataataata ataataatag aaaataaagt tgtcttcatg 2880
aaaaatgagg aaagagattg ctggggtgag aaacattaag atcaatgggc atatggtgac 2940
cttctatgcc ctagaaactc ttttanggta ttttctcctg gtatctcttt tacncatcgt 3000
tctatctgga aaaataggtg gatgagtgag ataataacgg tatatacttt ttaaaggtct 3060
aattgacata tataaattgc aagtatttca gatgtcaatt tgctaacctt gacacacata 3120
gacacacatg aaaacatcac cacattaata caatgtatgt atccatcatt ccaaaagctt 3180
ccctgtgtat ctttgtaact ctttcttcct ccctccactc cttgtcctct cgttcccaag 3240
aaaacattga tctgcttcct gtgaatataa attaacttac atttttaga gctttatata 3300
agtatgttct ctttactgtt tgtcttcctt cgctgcacag ttattttgag attcttcaag 3360
ttttttcttt atatcgatac ttcattcaca agaatatatt ttaattctag actatgtcac 3420
attgactttg tcgtctgcta aatccttagt gctcagatga cttgttcagg actctccttg 3480
aacctgtacc tctgttanat tgaaacttgt ctctactgtc tttttatttc aaacacagct 3540
tattaggtgt ctctcaaccc atcaaacnca caatctgagt ctttaggaga ttgctttgaa 3600
tttgtgctat tgacttatat ntatatnaaa tntgtaaatg tttggtaaaa atatcatcat 3660
gtacntttte ataattacgc tatntncaca tgatatatgt cagactctgg aaatatgcat 3720
gccacagaca cgtgtttctt gcctaaaggg gctgatggaa gacncacata cnaatagacg 3780
attgcagtag aatgagagtg gtggtctaan cagtacatgt cctgatgttg ctcggacagt 3840
tactacncca agagtacccc ctgcattgtc agggttagca tctcctggaa gcctcatgta 3900
aatgaagaat ttcatgctcc atccaggacc taatgaataa gaatctgcat tttagcaaga 3960
ccctcatatg attcatatac actttttttt ttttttttta gatggagtct cactcttgtc 4020
gcccaggctg gagtgcaatg gcatgatctt ggctcactgc aacctctgcc tcccgggttc 4080
aagtgattct cctgtctcag cctccctagt agctgggact acaggtgcat gccacagtgg 4140
ctggctaatt tttgtatttt tagtagagac agggtttcac cattttggtc aggctggtct 4200
tgaactcatg acctccggtg attccccccgc ctcggcttcc caaagtgctg ggattacaga 4260
catgagccac cacacccgcc ttattcgtat acncatttaa ttctgagaag cactctatag 4320
aaaataagaa taagaaaata ttgggctcac aggtgacatt aataagtaac tttatcgagt 4380
accccaaatt ttacctatgt ttggaagatg gggttaaaag gacacattga aaacaagaac 4440
tcattgtggc ttttttttcc tccttttga acagttttct atttctggaa tgttgtcaat 4500
tatatctgaa aggagaaatg caacatatct ggtgagttgc ccgtttctgt ctttgtccat 4560
ccttgaaaag ataagaagaa cagagtttta agagtcttaa gggaaacaca tctttgtctc 4620
ctatattact tgtgaatgtg gatatatgat tttgtttcaa tctattttgt gtcctaaggc 4680
tttttgcaac agaagttgga tatatcatta gaaacataaa ttgtaccatt taacatacat 4740
gaagtttatg tttaccttga cgttcttcta aaaagtgtcc tacaccggca ttgtccttgt 4800
aggcatattc acatgatcaa ataaaataat tagttttcaa ttaaggagaa tatttgagga 4860
aagaccgtac gtgttcatgt ggttcctgaa ggcagtccag tgagaaagta atatatgctt 4920
cattaaacaa tgcggacatt ttcagggttt ccctttttaa ccaaaatttg gaagcaatgt 4980
ggaatttact ggatgcatcc agccctgaaa tgaagatagg tttattgaat gtgccagcaa 5040
gtgcaggccc aggtctgagt gttcttcatt attatcaggt gagaggaagc ctgggagcaa 5100
acactgccag cagcatagct gggggaacgg gaattaccat cctgatcatc aacctgaaga 5160
agagcttggc ctatatccac atccacagtt gccagaaatt ttttgagacc aagtgcttta 5220
tggcttcctt ttccactgta tgtatttttt tttgtgtggg aagactaaga ttctgggtcc 5280
taatgtaagt aagaagccct cttctcctgt tccatgaaca ccatcctttt ctgtaacttc 5340
tattacacag tatagtggtt ctgtaagttc acacagccca gggagatgct ggctgcccac 5400
tcccctcaac ccaggcaaat tcctcggggt taaagttatc tactgcaagt gacgatctct 5460
gggtttttct gtgcctgtgt ttgtgtgtgt gtgtgtgtgt gtgtgtgtgt gtatgtgtca 5520
ctttaaaagg actggtcaga tggtagggag atgaaaacag gagatgctat aagaaaataa 5580
acttttgggg cgaataccaa tgtgactctt tttgtttgtc atttgttgct gttcaatagg 5640
aaattgtagt gatgatgctg tttctcacca ttctgggact tggtagtgct gtgtcactca 5700

caatctgtgg agctggggaa gaactcaaag gaaacaaggt agatagaagc ccgatataaa 5760
atcttgaatg acaggttaac gaattggagc tttattcctt aaaatatggc ctgggttttc 5820
tgaaacattt cttccagaaa atagtttctc caagttttat tactttggtt tacaaatctc 5880
acatttaaat cacattttat accataagta gcacacattt cataatattc ctctgaatga 5940
gggttgggat aataggactg atatgttaga aatgccttaa agtgtgtgga gcatgagaga 6000
tggatgtaca gaaggcttgt gaggaaacca cccaggtatc tggccttgtt ttctgcccca 6060
gaactagccg cctattcctg tttctgtttt attcctttgt ttcttgactt ttcctttcca 6120
acttgctcta aaacctcagt tttctttcct ttctgattca tgactaccaa atgttttcac 6180
ttgcctcacc cgtccattac acctttgata agaaccacca gaccttgtgc tcatgtactt 6240
gcccatgtct gatggaagaa acatactctc tccatctgtc cactttcctg aggcattcaa 6300
gtctagccac cttttaaaat cactctcctc caggctgggc acggtgtcac gcctgtaatc 6360
tcagcacttt gtgaggctga ggagggcgga tcacttgaag tcaggagttc aaaaccagcc 6420
tggccaaatg gcaaaaccaa atcttcttca attataacca aatcttaaac caaatctcta 6480
ctaaaaaata caacaaaaca aaacaacaac aacaaaaaca gaaaaggaaa cattagccca 6540
gcgtggtggc aggtacctga ggttccagat acttgggagg ctgaagcagg agaatcgctt 6600
gagcccaaga gatggaggtt gcagtgagcc gagatcatgc cactgcacca cagccagggt 6660
gacagagcca tacttcccag cacattggga ggccaaagct gaagaataat ttgaggtgag 6720
gatttggaga ccagcctggc caacatggtg aaactccgtc tgtactaaaa atataaaact 6780
tagtggggca tgggggcaca cacctgtaat ttcagctact taggaggctg aggcaggaga 6840
attgcttgaa cccgggaggc ggaagttgca gtgagccaag atcgtggcca ctgcactcca 6900
gcctgggtga catagtgaga ttctgtctca aaaaaaataa aagaaattta aaaaatcact 6960
ctcttccaaa gatagataaa taagacagca gatatactaa ggaataacct caccaacttg 7020
tcattgactg acatgatttc ttttggccca cttggccagc tagtctggtt tggttttctg 7080
gaaatgaaag aaataatcag agtttaatga cagagagcgt gagacccaga aagacaaaag 7140
tagatgaggt aagtctcttg agcgagactt ctagggatgg gaaatttgtg gtgattgata 7200
tgaaatgatt tttcccttat caggttccag aggatcgtgt ttatgaagaa ttaaacatat 7260
attcagctac ttacagtgag ttggaagacc caggggaaat gtctcctccc attgatttat 7320
aagaatcacg tgtccagaac actctgattc acagccaagg atccagaagg ccaaggtttt 7380
gttaaggggc tactggaaaa atttctattc tctccacagc ctgctggttt tacattagat 7440
ttattcgcct gataagaata ttttgtttct gctgcttctg tccaccttaa tatgctcctt 7500
ctatttgtag atatgataga ctcctatttt tcttgtttta tattatgacc acacacatct 7560
ctgctggaaa gtcaacatgt agtaagcaag atttaactgt ttgattataa ctgtgcaaat 7620
acagaaaaaa agaaggctgg ctgaaagttg agttaaactt tgacagtttg ataatatttg 7680
gttcttaggg ttttttttct ttttagcatt cttaatagtt acagttgggc atgatttgta 7740
ccatccaccc atacccacac agtcacagtc acacacacat atgtattact tacactatat 7800
ataacttcct atgcaaatat tttaccacca gtcaataata cattttgcc aagacatgaa 7860
gttttataaa gatctgtata attgcctgaa tcaccagcac attcactgac atgatattat 7920
ttgcagattg acaagtagga agtggggaac ttttattaag ttactcgttg tctggggagg 7980
taaataggtt aaaaacaggg aaattataag tgcagagatt aacatttcac aaatgtttag 8040
tgaaacattt gtgaaaaaag aagactaaat taagacctga gctgaaataa agtgacgtgg 8100
aaatggaaat aatggttata tctaaaacat gtagaaaaag agtaactggt agattttgtt 8160
aacaaattaa agaataaagt tagacaagca actggttgac taatacatta agcgtttgag 8220
tctaagatga aaggagaaca ctggttatgt tgatagaatg ataaaaaggg tcgggcgcgg 8280
aggctcacgc ctgtaatccc agccctttgg gaggccgagg tgggcagatc acgaagtcag 8340
tagtttgaga ccagcctggc caacatagtg aaaccccgtc tctactaaaa atacaaaaaa 8400
aaaattagct gggtgtggtg gcagtcacct gtagtcccag ctacttggga ggatgaggca 8460
ggagaatcgc ttgaacctgg gaggcggagg ttgcagtgag ccgagatcgc accagtgcac 8520
tccagccttg gtgacaatgg gagactccat ctcaaaaaaa aaaaaaaaaa aaaaaagata 8580
aaaagtcaga aatctgaaaa gtggaggaag agtacaaata gacctaaatt aagtctcatt 8640
ttttggcttt gattttgggg agacaaaggg aaatgcagcc atagagggcc tgatgacatc 8700
caatacatga gttctggtaa agataaaatt tgatacacgg tttggtgtca ttataagaga 8760
aatcattatt aaatgaagca agttaacact ctaagagaat tattttgaga tagaagtgaa 8820
gctaagctaa acttcacatg cctataattg gagggaaaaa ctaaggataa aatctagcct 8880
agaagataca ataattagtc ataaacatgc attgtgaaac tgtagagagc aggtagccca 8940
aaatagagaa agattagata aagagaaaat aagtatccat cagagacagt atctctaggc 9000
ttgggcaaga gaaaagtcca cagtgataag caactccacc taaggcatga atatgcggca 9060
gagaaaacag caatagtgaa tgaatgcaaa aggtgctgag caaattccac acatgagtat 9120
tgtgcatgag taaatgaata aaacatttgc aaagaccttt agagaaagag aatgggagca 9180
tatgtgcgaa ataagatagt tgattatgaa tagaaggtag tgaagaaaag caagctaaga 9240
aaaaattctg tttataaaag aaggaaaaga tagtttatgt ttttagccta agtataagag 9300
tcctacagat ggactgaaaa aaatcagtct gagagtatta gtcacaatta atgaaataat 9360
tacattttat gtattgagga tgccaagatt aaaaggtgac aggtagatgt taatttccct 9420
agattgtgaa agtgatcacg acaatcacac aacaaataat taagtgactt ggtatgcttt 9480
atttaattgt agggcctgag gttttccatt ctcatttttc taaaatacaa ttttgtttct 9540

ccaaatttga cagcagaata aaaaccctac cctttcactg tgtatcatgc taagctgcat 9600
ctctactctt gatcatctgt aggtattaat cacatcactt ccatggcatg gatgttcaca 9660
tacagactct taaccctggt ttaccaggac ctctaggagt ggatccaatc tatatcttta 9720
cagttgtata gtatatgata tctcttttat ttcactcaat ttatatttlc atcattgact 9780
acatatttct tatacacaac acacaattta tgaatttttt ctcaagatca ttctgagagt 9840
tgccccaccc tacctgcctt ttatagtacg cccacctcag gcagacacag agcacaatgc 9900
tggggttctc ttcacactat cactgcccca aattgtcttt ctaaatttca acttcaatgt 9960
catcttctcc atgaagacca ctgaatgaac acctttttcat ccagccttaa tttcttgctc 10020
cataactact ctatcccacg atgcagtatt gtatcattaa ttattagtgt gcttgtgacc 10080
tccttatgta ttctcaatta cctgtatttg tgcaataaat tggaataatg taacttgatt 10140
tcttatctgt gtttgtgttg gcatgcaaga tttaggtact tatcaagata atggggaatt 10200
aaggcatcaa taaaatgatg ccaaagacca agagcagttt ctgaagtcct ccttttcatc 10260
agctctttat caaacagaac actctataaa caacccatag ccagaaaaca ggatgtagga 10320
acaatcacca gcacactcta taaacaaccc atagccagaa aacagaatgt aaggacaatc 10380
accagccatc ttttgtcaat aattgatgga atagagttga aaggaactgg agcatgagtc 10440
atatttgacc agtcagtcct cactcttatt tacttgctat gtaaacttga gaaagctttt 10500
ttctctttgt gaacctcagg ttttacatct gaaaatgaga aatttggaac aaaagattcc 10560
taactggtct ttctgttccc atattctgtg atttttcaat atttaggatt tttggtaatc 10620
acaattactt agtttgtggt tgagatagca acacgaatca gaactatttg gtggacatat 10680
tttcaaagga gtagctctcc actttgggta aagaagtgat gcnggtcgtg gtggctcacg 10740
cctgtaatcc cagcacttta gggaggccaa ggcgggtgga tcacgaggtc aggagatcga 10800
gaccatcctg gctaacacgg tgaaaccccg tctctactaa aaaatacaaa aaattagcca 10860
ggcgtggtgg cgggcgcctg tagtcccacg tactcgggag gctgaggcag gagaatggca 10920
tgaaccaggg aggcggagct tgccgtgagc cgagatagcg ccactgcagt ccctcctggg 10980
caaaagagca agactgcgtc tcaaaaaaaa aaaaaaaaaa aaaaaaagaa gtgtgtggag 11040
tagcaggaca cctgcaacaa taatattttt ctaaatccct ctgaaaaatg ctaatcaaag 11100
ggttttttttc ctaaaaattg tcttagaaat aaaatttccc ctttgggaga ccgaggctgg 11160
cagatcacga ggtcaggaga tagagaccac ggtgaaaccc cgtctctact aaaaatacta 11220
aaaattagcc ggggngtggt ggtgggtaca cctgtagtcc cagctacttg gaggctgagg 11280
ctggagaatc acgtgaac 11298

<210> 2
<211> 735
<212> DNA
<213> Homo sapiens

<400> 2
atggacacag aaagtaatag gagagcaaat cttgctctcc cacaggagcc ttccagtgtg 60
cctgcatttg aagtcttgga aatatctccc caggaagtat cttcaggcag actattgaag 120
tcggcctcat ccccaccact gcatacatgg ctgacagttt tgaaaaaaga gcaggagttc 180
ctgggggtaa cacaaattct gactgctatg atatgccttt gttttggaac agttgtctgc 240
tctgtacttg atatttcaca cattgaggga gacatttttt catcatttaa agcaggttat 300
ccattctggg gagccatatt tttttctatt tctggaatgt tgtcaattat atctgaaagg 360
agaaatgcaa catatctggt gagaggaagc ctgggagcaa acactgccag cagcatagct 420
gggggaacgg gaattaccat cctgatcatc aacctgaaga agagcttggc ctatatccac 480
atccacagtt gccagaaatt ttttgagacc aagtgcttta tggcttcctt ttccactgaa 540
attgtagtga tgatgctgtt tctcaccatt ctgggacttg gtagtgctgt gtcactcaca 600
atctgtggag ctggggaaga actcaaagga aacaaggttc cagaggatcg tgtttatgaa 660
gaattaaaca tatattcagc tacttacagt gagttggaag acccagggga aatgtctcct 720
cccattgatt tataa 735

<210> 3
<211> 244
<212> PRT
<213> Homo sapiens

<400> 3
Met Asp Thr Glu Ser Asn Arg Arg Ala Asn Leu Ala Leu Pro Gln Glu
1 5 10 15

Pro Ser Ser Val Pro Ala Phe Glu Val Leu Glu Ile Ser Pro Gln Glu
20 25 30

Val Ser Ser Gly Arg Leu Leu Lys Ser Ala Ser Ser Pro Pro Leu His
35 40 45

Thr Trp Leu Thr Val Leu Lys Lys Glu Gln Glu Phe Leu Gly Val Thr
50 55 60

Gln Ile Leu Thr Ala Met Ile Cys Leu Cys Phe Gly Thr Val Val Cys
65 70 75 80

Ser Val Leu Asp Ile Ser His Ile Glu Gly Asp Ile Phe Ser Ser Phe
85 90 95

Lys Ala Gly Tyr Pro Phe Trp Gly Ala Ile Phe Phe Ser Ile Ser Gly
100 105 110

Met Leu Ser Ile Ile Ser Glu Arg Arg Asn Ala Thr Tyr Leu Val Arg
115 120 125

Gly Ser Leu Gly Ala Asn Thr Ala Ser Ser Ile Ala Gly Gly Thr Gly
130 135 140

Ile Thr Ile Leu Ile Ile Asn Leu Lys Lys Ser Leu Ala Tyr Ile His
145 150 155 160

Ile His Ser Cys Gln Lys Phe Phe Glu Thr Lys Cys Phe Met Ala Ser
165 170 175

Phe Ser Thr Glu Ile Val Val Met Met Leu Phe Leu Thr Ile Leu Gly
180 185 190

Leu Gly Ser Ala Val Ser Leu Thr Ile Cys Gly Ala Gly Glu Glu Leu
195 200 205

Lys Gly Asn Lys Val Pro Glu Asp Arg Val Tyr Glu Glu Leu Asn Ile
210 215 220

Tyr Ser Ala Thr Tyr Ser Glu Leu Glu Asp Pro Gly Glu Met Ser Pro
225 230 235 240

Pro Ile Asp Leu

<210> 4
<211> 15
<212> DNA
<213> Homo sapiens

<400> 4
atcacaagta aaagc 15

<210> 5
<211> 15
<212> DNA
<213> Homo sapiens

<400> 5
atcacaaata aaagc 15

<210> 6
<211> 15

<212> DNA
<213> Homo sapiens

<400> 6
cctcatcccc accac 15

<210> 7
<211> 15
<212> DNA
<213> Homo sapiens

<400> 7
cctcatctcc accac 15

<210> 8
<211> 15
<212> DNA
<213> Homo sapiens

<400> 8
gtattaagat gatat 15

<210> 9
<211> 15
<212> DNA
<213> Homo sapiens

<400> 9
gtattaacat gatat 15

<210> 10
<211> 15
<212> DNA
<213> Homo sapiens

<400> 10
tataacatag atatg 15

<210> 11
<211> 15
<212> DNA
<213> Homo sapiens

<400> 11
tataacacag atatg 15

<210> 12
<211> 15
<212> DNA
<213> Homo sapiens

<400> 12
gggggaacgg gaatt 15

<210> 13
<211> 15
<212> DNA

<213> Homo sapiens

<400> 13
ggggggaatgg gaatt 15

<210> 14
<211> 15
<212> DNA
<213> Homo sapiens

<400> 14
ttggcctata tccac 15

<210> 15
<211> 15
<212> DNA
<213> Homo sapiens

<400> 15
ttggcctgta tccac 15

<210> 16
<211> 15
<212> DNA
<213> Homo sapiens

<400> 16
ttccttttcc actgt 15

<210> 17
<211> 15
<212> DNA
<213> Homo sapiens

<400> 17
ttcctttgcc actgt 15

<210> 18
<211> 15
<212> DNA
<213> Homo sapiens

<400> 18
ttttttttttg tgtgg 15

<210> 19
<211> 15
<212> DNA
<213> Homo sapiens

<400> 19
ttttttttgtg tgtgg 15

<210> 20
<211> 15
<212> DNA
<213> Homo sapiens

<400> 20
tttttgtgtg ggaag 15

<210> 21
<211> 15
<212> DNA
<213> Homo sapiens

<400> 21
tttttgtatg ggaag 15

<210> 22
<211> 15
<212> DNA
<213> Homo sapiens

<400> 22
ggaaaaattt ctatt 15

<210> 23
<211> 15
<212> DNA
<213> Homo sapiens

<400> 23
ggaaaaactt ctatt 15

<210> 24
<211> 15
<212> DNA
<213> Homo sapiens

<400> 24
atttattcgc ctgat 15

<210> 25
<211> 15
<212> DNA
<213> Homo sapiens

<400> 25
atttatttgc ctgat 15

<210> 26
<211> 15
<212> DNA
<213> Homo sapiens

<400> 26
ctattcatca caagt 15

<210> 27
<211> 15
<212> DNA
<213> Homo sapiens

<400> 27
caacaggctt ttact 15

<210> 28
<211> 15
<212> DNA
<213> Homo sapiens

<400> 28
ctattcatca caaat 15

<210> 29
<211> 15
<212> DNA
<213> Homo sapiens

<400> 29
caacaggctt ttatt 15

<210> 30
<211> 15
<212> DNA
<213> Homo sapiens

<400> 30
agtcggcctc atccc 15

<210> 31
<211> 15
<212> DNA
<213> Homo sapiens

<400> 31
tatgcagtgg tgggg 15

<210> 32
<211> 15
<212> DNA
<213> Homo sapiens

<400> 32
agtcggcctc atctc 15

<210> 33
<211> 15
<212> DNA
<213> Homo sapiens

<400> 33
tatgcagtgg tggag 15

<210> 34
<211> 15
<212> DNA
<213> Homo sapiens

<400> 34

ttcccagtat taaga 15

<210> 35
<211> 15
<212> DNA
<213> Homo sapiens

<400> 35
ttataaatat catct 15

<210> 36
<211> 15
<212> DNA
<213> Homo sapiens

<400> 36
ttcccagtat taaca 15

<210> 37
<211> 15
<212> DNA
<213> Homo sapiens

<400> 37
ttataaatat catgt 15

<210> 38
<211> 15
<212> DNA
<213> Homo sapiens

<400> 38
agcatatata acata 15

<210> 39
<211> 15
<212> DNA
<213> Homo sapiens

<400> 39
aatgagcata tctat 15

<210> 40
<211> 15
<212> DNA
<213> Homo sapiens

<400> 40
agcatatata acaca 15

<210> 41
<211> 15
<212> DNA
<213> Homo sapiens

<400> 41
aatgagcata tctgt 15

<210> 42
<211> 15
<212> DNA
<213> Homo sapiens

<400> 42
atagctgggg gaacg 15

<210> 43
<211> 15
<212> DNA
<213> Homo sapiens

<400> 43
gatggtaatt cccgt 15

<210> 44
<211> 15
<212> DNA
<213> Homo sapiens

<400> 44
atagctgggg gaatg 15

<210> 45
<211> 15
<212> DNA
<213> Homo sapiens

<400> 45
gatggtaatt cccat 15

<210> 46
<211> 15
<212> DNA
<213> Homo sapiens

<400> 46
aagagcttgg cctat 15

<210> 47
<211> 15
<212> DNA
<213> Homo sapiens

<400> 47
gtggatgtgg atata 15

<210> 48
<211> 15
<212> DNA
<213> Homo sapiens

<400> 48
aagagcttgg cctgt 15

<210> 49
<211> 15
<212> DNA
<213> Homo sapiens

<400> 49
gtggatgtgg ataca 15

<210> 50
<211> 15
<212> DNA
<213> Homo sapiens

<400> 50
tatggcttcc ttttc 15

<210> 51
<211> 15
<212> DNA
<213> Homo sapiens

<400> 51
atacatacag tggaa 15

<210> 52
<211> 15
<212> DNA
<213> Homo sapiens

<400> 52
tatggcttcc tttgc 15

<210> 53
<211> 15
<212> DNA
<213> Homo sapiens

<400> 53
atacatacag tggca 15

<210> 54
<211> 15
<212> DNA
<213> Homo sapiens

<400> 54
tatgtatttt ttttt 15

<210> 55
<211> 15
<212> DNA
<213> Homo sapiens

<400> 55
gtcttcccac acaaa 15

<210> 56
<211> 15
<212> DNA
<213> Homo sapiens

<400> 56
tatgtatttt tttgt 15

<210> 57
<211> 15
<212> DNA
<213> Homo sapiens

<400> 57
gtcttcccac acaca 15

<210> 58
<211> 15
<212> DNA
<213> Homo sapiens

<400> 58
tatttttttt tgtgt 15

<210> 59
<211> 15
<212> DNA
<213> Homo sapiens

<400> 59
cttagtcttc ccaca 15

<210> 60
<211> 15
<212> DNA
<213> Homo sapiens

<400> 60
tatttttttt tgtat 15

<210> 61
<211> 15
<212> DNA
<213> Homo sapiens

<400> 61
cttagtcttc ccata 15

<210> 62
<211> 15
<212> DNA
<213> Homo sapiens

<400> 62
gctactggaa aaatt 15

<210> 63

<211> 15
<212> DNA
<213> Homo sapiens

<400> 63
ggagagaata gaaat 15

<210> 64
<211> 15
<212> DNA
<213> Homo sapiens

<400> 64
gctactggaa aaact 15

<210> 65
<211> 15
<212> DNA
<213> Homo sapiens

<400> 65
ggagagaata gaagt 15

<210> 66
<211> 15
<212> DNA
<213> Homo sapiens

<400> 66
cattagattt attcg 15

<210> 67
<211> 15
<212> DNA
<213> Homo sapiens

<400> 67
attcttatca ggcga 15

<210> 68
<211> 15
<212> DNA
<213> Homo sapiens

<400> 68
cattagattt atttg 15

<210> 69
<211> 15
<212> DNA
<213> Homo sapiens

<400> 69
attcttatca ggcaa 15

<210> 70
<211> 10

<212> DNA
<213> Homo sapiens

<400> 70
ttcatcacaa 10

<210> 71
<211> 10
<212> DNA
<213> Homo sapiens

<400> 71
caggctttta 10

<210> 72
<211> 10
<212> DNA
<213> Homo sapiens

<400> 72
cggcctcatc 10

<210> 73
<211> 10
<212> DNA
<213> Homo sapiens

<400> 73
gcagtggtgg 10

<210> 74
<211> 10
<212> DNA
<213> Homo sapiens

<400> 74
ccagtattaa 10

<210> 75
<211> 10
<212> DNA
<213> Homo sapiens

<400> 75
taaatatcat 10

<210> 76
<211> 10
<212> DNA
<213> Homo sapiens

<400> 76
atatataaca 10

<210> 77
<211> 10
<212> DNA

<213> Homo sapiens

<400> 77
gagcatatct 10

<210> 78
<211> 10
<212> DNA
<213> Homo sapiens

<400> 78
gctgggggaa 10

<210> 79
<211> 10
<212> DNA
<213> Homo sapiens

<400> 79
ggtaattccc 10

<210> 80
<211> 10
<212> DNA
<213> Homo sapiens

<400> 80
agcttggcct 10

<210> 81
<211> 10
<212> DNA
<213> Homo sapiens

<400> 81
gatgtggata 10

<210> 82
<211> 10
<212> DNA
<213> Homo sapiens

<400> 82
ggcttccttt 10

<210> 83
<211> 10
<212> DNA
<213> Homo sapiens

<400> 83
catacagtgg 10

<210> 84
<211> 10
<212> DNA
<213> Homo sapiens

<400> 84
gtatttttt 10

<210> 85
<211> 10
<212> DNA
<213> Homo sapiens

<400> 85
ttcccacaca 10

<210> 86
<211> 10
<212> DNA
<213> Homo sapiens

<400> 86
ttttttttgt 10

<210> 87
<211> 10
<212> DNA
<213> Homo sapiens

<400> 87
agtcttccca 10

<210> 88
<211> 10
<212> DNA
<213> Homo sapiens

<400> 88
actggaaaaa 10

<210> 89
<211> 10
<212> DNA
<213> Homo sapiens

<400> 89
gagaatagaa 10

<210> 90
<211> 10
<212> DNA
<213> Homo sapiens

<400> 90
tagatttatt 10

<210> 91
<211> 10
<212> DNA
<213> Homo sapiens

<400> 91
cttatcaggc 10

<210> 92
<211> 22
<212> DNA
<213> Homo sapiens

<400> 92
ggtgcaattg gataacttct gc 22

<210> 93
<211> 24
<212> DNA
<213> Homo sapiens

<400> 93
ctaagcaccg tgactatgac ttcc 24

<210> 94
<211> 23
<212> DNA
<213> Homo sapiens

<400> 94
gtggggacaa ttccagaaga agg 23

<210> 95
<211> 25
<212> DNA
<213> Homo sapiens

<400> 95
cctgtctcat acccagaaaa tgagc 25

<210> 96
<211> 26
<212> DNA
<213> Homo sapiens

<400> 96
cccatctagc tattcaaagc cattcc 26

<210> 97
<211> 22
<212> DNA
<213> Homo sapiens

<400> 97
tgttggccat tctcgtccta gc 22

<210> 98
<211> 24
<212> DNA
<213> Homo sapiens

<400> 98

agtattggcc ttccatgcat tagg 24

<210> 99
<211> 26
<212> DNA
<213> Homo sapiens

<400> 99
tgctagaagt atgttcctgg agttgg 26

<210> 100
<211> 22
<212> DNA
<213> Homo sapiens

<400> 100
atgtggttcc tgaaggcagt cc 22

<210> 101
<211> 23
<212> DNA
<213> Homo sapiens

<400> 101
aaaacccaga gatcgtcact tgc 23

<210> 102
<211> 25
<212> DNA
<213> Homo sapiens

<400> 102
cctcggggtt aaagttatct actgc 25

<210> 103
<211> 24
<212> DNA
<213> Homo sapiens

<400> 103
tcctcacaag ccttctgtac atcc 24

<210> 104
<211> 26
<212> DNA
<213> Homo sapiens

<400> 104
acccagaaag acaaaagtag atgagg 26

<210> 105
<211> 23
<212> DNA
<213> Homo sapiens

<400> 105
ttccagcaga gatgtgtgtg gtc 23

<210> 106
<211> 20
<212> DNA
<213> Homo sapiens

<400> 106
ttgggacaca agtggggaca 20

<210> 107
<211> 20
<212> DNA
<213> Homo sapiens

<400> 107
gaaggctcct gtgggagagc 20

<210> 108
<211> 21
<212> DNA
<213> Homo sapiens

<400> 108
tcctgctagt ctcaggcaaa a 21

<210> 109
<211> 21
<212> DNA
<213> Homo sapiens

<400> 109
ctccagggag agagaaccac a 21

<210> 110
<211> 20
<212> DNA
<213> Homo sapiens

<400> 110
ttggcatgga ttttgctcca 20

<210> 111
<211> 21
<212> DNA
<213> Homo sapiens

<400> 111
tgcttctggc tcttcccaaa a 21

<210> 112
<211> 22
<212> DNA
<213> Homo sapiens

<400> 112
cgtgtggatc atttctcagg ac 22

<210> 113
<211> 20
<212> DNA
<213> Homo sapiens

<400> 113
tcaagtggcc cttaggcaag 20

<210> 114
<211> 21
<212> DNA
<213> Homo sapiens

<400> 114
tgcggacatt ttcagggttt c 21

<210> 115
<211> 20
<212> DNA
<213> Homo sapiens

<400> 115
tctccctggg ctgtgtgaac 20

<210> 116
<211> 21
<212> DNA
<213> Homo sapiens

<400> 116
ctgggttttt ctgtgcctgt g 21

<210> 117
<211> 21
<212> DNA
<213> Homo sapiens

<400> 117
tcccaaccct cattcagagg a 21

<210> 118
<211> 18
<212> DNA
<213> Homo sapiens

<400> 118
aagtctcttg agcgagac 18

<210> 119
<211> 18
<212> DNA
<213> Homo sapiens

<400> 119
aaggtggaca gaagcagc 18

<210> 120
<211> 23
<212> DNA
<213> Homo sapiens

<400> 120
ggcaaaatta tgctccagga gtc 23

<210> 121
<211> 23
<212> DNA
<213> Homo sapiens

<400> 121
aagcaccgtg actatgactt ccc 23

<210> 122
<211> 24
<212> DNA
<213> Homo sapiens

<400> 122
ttggtcagtt acttggatgc tctg 24

<210> 123
<211> 20
<212> DNA
<213> Homo sapiens

<400> 123
tggctcttcc caaaaggacc 20

<210> 124
<211> 24
<212> DNA
<213> Homo sapiens

<400> 124
cgtgtggatc atttctcagg acag 24

<210> 125
<211> 23
<212> DNA
<213> Homo sapiens

<400> 125
tctcatggga taaaactgtg ggc 23

<210> 126
<211> 24
<212> DNA
<213> Homo sapiens

<400> 126
tttacctatg tttggaagat gggg 24

<210> 127

<211> 24
<212> DNA
<213> Homo sapiens

<400> 127
gaatatgcct acaaggacaa tgcc 24

<210> 128
<211> 23
<212> DNA
<213> Homo sapiens

<400> 128
tccagccctg aaatgaagat agg 23

<210> 129
<211> 22
<212> DNA
<213> Homo sapiens

<400> 129
actttaaccc cgaggaattt gc 22

<210> 130
<211> 23
<212> DNA
<213> Homo sapiens

<400> 130
aaaggactgg tcagatggta ggg 23

<210> 131
<211> 24
<212> DNA
<213> Homo sapiens

<400> 131
tctcatgctc cacacacttt aagg 24

<210> 132
<211> 24
<212> DNA
<213> Homo sapiens

<400> 132
tcttgagcga gacttctagg gatg 24

<210> 133
<211> 24
<212> DNA
<213> Homo sapiens

<400> 133
catgttgact ttccagcaga gatg 24

<210> 134
<211> 24

<212> DNA
<213> Homo sapiens

<400> 134
catgaggtaa cccatttcaa ctgc 24

<210> 135
<211> 22
<212> DNA
<213> Homo sapiens

<400> 135
ccgtgactat gacttcccct gc 22

<210> 136
<211> 25
<212> DNA
<213> Homo sapiens

<400> 136
tggtcagtta cttggatgct ctgag 25

<210> 137
<211> 23
<212> DNA
<213> Homo sapiens

<400> 137
ttcccaaaag gacccagtta gtg 23

<210> 138
<211> 23
<212> DNA
<213> Homo sapiens

<400> 138
ttggaagatg gggttaaaag gac 23

<210> 139
<211> 20
<212> DNA
<213> Homo sapiens

<400> 139
tgcctacaag gacaatgccg 20

<210> 140
<211> 21
<212> DNA
<213> Homo sapiens

<400> 140
tgaatgtgcc agcaagtgca g 21

<210> 141
<211> 20
<212> DNA

<213> Homo sapiens

<400> 141
gaatttgcct gggttgaggg 20

<210> 142
<211> 25
<212> DNA
<213> Homo sapiens

<400> 142
ggtcagatgg tagggagatg aaaac 25

<210> 143
<211> 23
<212> DNA
<213> Homo sapiens

<400> 143
tcatgctcca cacactttaa ggc 23

<210> 144
<211> 22
<212> DNA
<213> Homo sapiens

<400> 144
tgagcgagac ttctagggat gg 22

<210> 145
<211> 24
<212> DNA
<213> Homo sapiens

<400> 145
gactttccag cagagatgtg tgtg 24

SCANNED, # 14

United States Patent & Trademark Office

Office of Initial Patent Examination -- Scanning Division

Application deficiencies found during scanning:

☒ Page(s)_____ of Specification were not present for scanning.
(Document title)

☐ Page(s)_____ of _____ were not present for scanning.
(Document title)

☐ ***Scanned copy is best available.***